UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Scienze Economiche "Marco Fanno"

VARIANCE CLUSTERING IMPROVED DYNAMIC CONDITIONAL CORRELATION MGARCH ESTIMATORS

GIAN PIERO AIELLI
University of Padova

MASSIMILIANO CAPORIN
University of Padova

May 2011

"MARCO FANNO" WORKING PAPER N.133

Variance Clustering Improved Dynamic Conditional Correlation MGARCH Estimators*

Gian Piero Aielli† Massimiliano Caporin‡

May 10, 2011

Abstract

It is well-known that the estimated GARCH dynamics exhibit common patterns. Starting from this fact we extend the Dynamic Conditional Correlation (DCC) model by allowing for a clustering structure of the univariate GARCH parameters. The model can be estimated in two steps, the first devoted to the clustering structure, and the second focusing on correlation parameters. Differently from the traditional two-step DCC estimation, we get large system feasibility of the joint estimation of the whole set of model's parameters. We also present a new approach to the clustering of GARCH processes, which embeds the asymptotic properties of the univariate quasi-maximum-likelihood GARCH estimators into a Gaussian mixture clustering algorithm. Unlike other GARCH clustering techniques, our method logically leads to the selection of the optimal number of clusters.

Keywords: *dynamic conditional correlations, time series clustering, multivariate GARCH, composite likelihood.* JEL codes: C32, C38, C53, C51, C52, C58.

*We thank Christian Brownlees, Giorgio Calzolari, Eduardo Otranto, Marc Paolella, Michael McAleer, Nicola Sartori and the participants to the Computational and Financial Econometrics 2011 conference in London for their helpful comments and suggestions.

†Department of Economics and Management *Marco Fanno*, University of Padova

‡Corresponding author: Department of Economics and Management *Marco Fanno*, University of Padova, Via del Santo 33, Padova, Italy, Tel: +39 049/8274258. Fax: +39 049/8274211. Email: `massimiliano.caporin@unipd.it`.

1 Introduction

The Dynamic Conditional Correlation GARCH (DCC) model (Engle, 2002) has recently become one of the most popular tools for the estimation of multivariate asset volatility dynamics (for applications see, e.g., Cappiello et al., 2006; Billio et al., 2006; Billio and Caporin, 2009; Franses and Hafner, 2009; Pesaran and Pesaran, 2007; some theoretical results are in Engle and Sheppard, 2001; McAleer et al., 2008; Aielli, 2006, 2011). The basic idea of the DCC modeling approach is to obtain first the variance process via simple univariate specifications, and, then to build a model for the correlation process from some appropriate function of the univariate variance standardized returns. As a direct advantage of such a modeling strategy, we obtain a two-step estimation procedure that is feasible with large systems: in the first step, the univariate variance processes are estimated one at a time, and, then, in the second step, the correlation process is estimated from the estimated standardized returns provided by the first step. Under appropriate specifications of the correlation process such as the cDCC specification adopted in this paper (Aielli, 2006, 2011), the two-step estimation procedure is shown to have desirable theoretical and empirical properties.

Most proposed extensions of the DCC model have been developed with the aim of providing more flexibility to the model of the correlation process. Two examples are the Asymmetric DCC model by Cappiello et al. (2006), and the Flexible DCC model by Billio et al. (2006); see also Billio and Caporin (2009), Engle and Kelly (2009) and Franses and Hafner (2009) for other applications. In particular, with the Flexible DCC model a rather rich parametrization of the correlation process is obtained by relying on *a priori* knowledge of the presence of some asset partition. The focus of this paper is different: we aim at improving the flexibility of the DCC modeling approach by focusing attention on the "variance side," rather than on the "correlation side," of the DCC model. Our final interest is to avoid, or reduce, the loss of efficiency (if any) that potentially affects the traditional two-step DCC estimator; in fact, in that case, the variance processes are estimated one at a time and independently from the information embedded in the correlation dynamics. With this aim in mind, we take our cue from the well-known fact that the univariate estimates of GARCH dynamic parameters are often similar. We then allow the specification of the cDCC model to take advantage of this empirical finding by assuming that there can be assets in the selected asset sample that share common GARCH dynamics. We call this assumption *variance clustering* (VC) *assumption*, and the resulting model, VC-cDCC model. If the number of clusters is equal to the number of assets, we obtained the cDCC model as a special case of the VC-cDCC model. Given an estimate of the underlying asset variance clustering structure, the joint estimation of the whole set of model dynamic parameters — including the variance dynamic parameters and the correlation dynamic parameters — becomes possible.

Under a fully Bayesian framework, Bauwens and Rombouts (2007) recently addressed the estimation of the underlying GARCH dynamic clustering structure. The refined clustering approach proposed by the authors makes heavy use of Monte Carlo Markov Chain (MCMC) methods. As a result, their approach is computationally intensive, in particular with a large number of assets. A mixed frequentist-bayesian approach is proposed by Brownless (2010), in which the parameters of a set of GARCH processes are modeled as functions of observed regressors and unobserved idiosyncratic shocks. On the contrary, in our paper we adopt a fully frequentist approach, with the explicit aim of providing a fast and friendly clustering methodology, in accordance with the motivations behind the DCC estimation paradigm, as originally proposed by Engle (2002). Our clustering algorithm constitutes a new contribution to the literature on the clustering of GARCH processes, or, in more general terms, on the clustering of financial time series (see Liao (2005) for a survey on time series clustering). We know from Elie and Jeantheau (1995) that the asymptotic distribution of the univariate Quasi Maximum Likelihood (QML) GARCH estimator is Gaussian. Therefore, a possible choice is to consider a Gaussian mixture (GM) clustering framework. Unlike Bauwens and Rombouts (2007), who specify a GM model for the observed series as a consequence of the assumption of Gaussian conditionally distributed returns, our GM model is specific to the sample of the univariate QML estimates of the GARCH dynamic parameters. We thus change the focus of the GM clustering from the return process to the GARCH estimators. As a result, our approach is more flexible since we are working on a distribution-free GARCH framework where asymptotic properties of the QMLE univariate GARCH estimators are valid anyway. We also point out that, to estimate the peculiar mixture component covariance matrix implied by our GM model specification, we resort to an appropriate combination of standard univariate QML estimation outputs. Furthermore — and differently from other *ad hoc*, non-Bayesian GARCH clustering techniques (e.g., Otranto, 2008) — our clustering methodology logically leads to a BIC-based selection procedure for the identification of the optimal number of clusters.

In order to deal with many assets, however, the reduction of the parameter dimensionality allowed by the variance clustering structure generally is not enough to allow the joint estimation of the variance/correlation dynamic parameters. The infeasibility is due to the fact that the dimensionality of the variance/correlation intercept parameters is $O(N^2)$, where N is the number of series. Following Aielli (2006), we successfully solve this problem by resorting to an *ad hoc* generalized profile QML estimator (Severini, 1998). We thus treat the variance/correlation intercept parameters as nuisance parameters, and the variance/correlation dynamic parameters as parameters of interest. We then replace the nuisance parameters in the model quasi-log-likelihood with an estimator that can be easily computed conditionally on the parameters of interest, and that is shown to possess some relevant properties. Thanks to such an estimation device, the joint estimation of the variance/correlation dynamic parameters becomes feasible (we call it the *joint* VC-cDCC estimator). The joint VC-cDCC estimator is shown to be superior to the traditional two-step cDCC estimator on the basis of both simulations and applications to real data.

For the joint VC-cDCC estimator to be feasible, it is required that the number of clusters is small, which, fortunately, is precisely what happens when dealing with large systems. An estimator of the VC-cDCC model that is feasible irrespective of the number of clusters is, however, also suggested in this paper (the *sequential* VC-cDCC estimator). With the sequential estimator, the asset variance dynamics are still estimated jointly, but separately from the correlation process. As objective functions for the variance dynamic parameters we adopt univariate composite likelihoods (Pakel et al., 2011). A special case of the sequential estimator is the traditional two-step cDCC estimator. The sequential estimator is extremely rapid to compute, and, surprisingly, does not cause any sensible loss of efficiency with respect to the joint estimator. From a theoretical perspective, this finding seems to suggest that the dynamic variance parameter and the dynamic correlation parameter are practically orthogonal (with respect to the adopted pseudo-likelihood).

The rest of the paper is organized as follows: section 2 introduces the VC-cDCC model; sections 3-4 describe the VC-cDCC estimator and discuss simulation results; section 5 reports some applications to real data, and, finally, section 6 concludes.

2 The VC-cDCC Model

Before introducing the VC-cDCC model we review the DCC modelling approach. Denote by $y_t = [y_{1,t}, \dots, y_{N,t}]'$ the $N \times 1$ vector of the asset returns at time t, and assume that

$$E_{t-1}[y_t] = 0\,, \;\; E_{t-1}[y_t y_t'] = H_t,$$

where $E_t[\,\cdot\,]$ is the conditional expectation on $y_t, y_{t-1}, \dots$ The asset conditional covariance matrix can be written as

$$H_t = D_t^{1/2} R_t \, D_t^{1/2}, \tag{1}$$

where $R_t = [\rho_{ij,t}]$ is the asset conditional correlation matrix, and $D_t = \mathrm{diag}(h_{1,t}, \ldots, h_{N,t})$ is the diagonal matrix of the asset conditional variances. By construction, R_t is the conditional covariance matrix of the asset standardized returns, that is, $E_{t-1}[\varepsilon_t \varepsilon_t'] = R_t$, where $\varepsilon_t = [\varepsilon_{1,t}, \ldots, \varepsilon_{N,t}]'$, and

$$\varepsilon_{i,t} = y_{i,t}/\sqrt{h_{i,t}}.$$

Engle (2002), following Bollerslev (1990), suggests modelling the right hand side of Eq. 1 rather than H_t directly. The resulting model is called the DCC model. In this work, we adopt the cDCC specification proposed by Aielli (2006, 2011), and then later used in empirical applications by Engle et al. (2009), Engle and Kelly (2009), and Hafner and Reznikova (2010).

2.1 The cDCC Model

The cDCC model assumes that the elements of D_t follow GARCH(P,Q) processes, Bollerslev (1986),

$$h_{i,t} = w_i + \sum_{p=1}^{P} a_{i,p}\, y_{i,t-p}^2 + \sum_{q=1}^{Q} b_{i,q}\, h_{i,t-q} \tag{2}$$

where the GARCH parameters, w_i, $a_{i,p}$, and $b_{i,q}$, are non-negative to ensure positivity of $h_{i,t}$. We further assume weakly stationarity of $y_{i,t}$, that is, parameters satisfy $\sum_{p=1}^{P} a_{i,p} + \sum_{q=1}^{Q} b_{i,q} < 1$. We note that, under stationarity, w_i can be reparametrized as

$$w_i = \left(1 - \sum_{p=1}^{P} a_{i,p} - \sum_{q=1}^{Q} b_{i,q}\right)\tau_i, \;\; \tau_i \geq 0, \tag{3}$$

and where it can be easily proved that

$$\tau_i = E[y_{i,t}^2] \tag{4}$$

and τ_i is the stationary second order moment of $y_{i,t}$. We then define the cDCC correlation process as

$$\rho_{ij,t} = \frac{(1-\alpha_{ij}-\beta_{i,j})\, s_{ij,t-1} + \alpha_{ij}\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \beta_{ij}\rho_{ij,t-1}}{\sqrt{\{(1-\alpha_{ii}-\beta_{ii})\, s_{ii,t-1} + \alpha_{ii}\varepsilon_{i,t-1}^2 + \beta_{ii}\rho_{ii,t-1}\}\{(1-\alpha_{jj}-\beta_{jj})\, s_{jj,t-1} + \alpha_{jj}\varepsilon_{j,t-1}^2 + \beta_{jj}\rho_{jj,t-1}\}}}, \tag{5}$$

where

$$s_{ij,t} = s_{ij}/\sqrt{q_{i,t}\, q_{j,t}} \tag{6}$$

and

$$q_{i,t} = (1-\alpha_{ii}-\beta_{ii}) + \alpha_{ii}\,\varepsilon_{i,t-1}^2\, q_{i,t-1} + \beta_{ii}\, q_{i,t-1}\,.$$

We collect the parameters of the correlation process in the matrices $A = [\alpha_{ij}]$, $B = [\beta_{ij}]$ and $S = [s_{ij}]$. The representation in (5) of $\rho_{ij,t}$ highlights that the cDCC model combines into a correlation-like ratio, a sort of GARCH device for both the relevant innovations and the past correlations (the numerator and the terms in braces at the denominators of eq (5), respectively). We also note that the denominator of $s_{ij,t}$ is an *ad hoc* correction required for tractability.[1] For $R_t = [\rho_{ij,t}]$ to be a proper correlation matrix[2] we must impose that A, B and $(\iota\iota' - A - B) \odot S$ are positive semi-definite, where $\odot$ denotes the element-wise (Hadamard) matrix product, and ι denotes the $N \times 1$ unit vector. To achieve identification, S must have unit diagonal elements $s_{ii} = 1$, $i = 1, 2, \ldots, N$. If $\alpha_{ij} = \alpha$ and $\beta_{ij} = \beta$ for $i \leq j = 1, 2, \ldots, N$, the model is called Scalar cDCC. In the last case, R_t is a proper correlation matrix provided that S is positive semi-definite, $\alpha \geq 0$, $\beta \geq 0$ and $\alpha + \beta \leq 1$. Setting $\alpha = \beta = 0$ yields the Constant Conditional Correlation (CCC) model of Bollerslev (1990), where $R_t = S$ at each point in time.

It can be shown that the cDCC model satisfies

$$S = E[\{Q_t^* \varepsilon_t\}\{Q_t^* \varepsilon_t\}'], \tag{7}$$

where $Q_t^* = \text{diag}(\sqrt{q_{1,t}}, \ldots, \sqrt{q_{N,t}})$. This property plays a crucial role in the construction of reasonable large system estimators (see section 4.1). For further details on the cDCC model properties, see Aielli (2011).

2.2 The VC-cDCC Model

As we state in the introduction, we assume that the N assets are grouped into K clusters. Assets within a given cluster share a common GARCH dynamic, whereas assets belonging to different clusters can exhibit different GARCH dynamics. Our purpose is to combine the cDCC dynamic with a restriction on the possible GARCH dynamics followed by the N assets. We stress that our modelling approach is different from several contributions of the MGARCH literature that focused on correlation dynamic parameters. If we introduce variance clustering into the cDCC model we obtain the Variance Clustering cDCC model (VC-cDCC). Denote as $C_i = [a_{i1}, a_{i2}, \ldots, a_{iP}, b_{i1}, b_{i2}, \ldots, b_{iQ}]'$ the $(P + Q) \times 1$ vector collecting the GARCH dynamic parameters of the i-th asset. We assume

[1] Compared with the traditional — and equivalent — representation of $\rho_{ij,t}$ in terms of rescaled elements of the correlation driving process, Q_t (see Engle, 2002, Aielli, 2006, 2011), the representation of $\rho_{ij,t}$ provided here is more immediate and clearly highlights the construction process of the correlation dynamic.

[2] A correlation matrix is a positive semi-definite matrix with unit diagonal elements.

that for any pair (C_i, C_j)

$$C_i = C_j \text{ if } \pi_i = \pi_j, \tag{8}$$

where $\pi_i \in \{1, 2, \ldots, K\}$, $i = 1, 2, \ldots, N$. The partition vector, $\pi = [\pi_1, \pi_2, \ldots, \pi_N]'$, arranges the N assets into the K clusters. The k-th cluster is defined as the subset of assets $\{y_{it} : \pi_i = k\}$. The size of the k-th cluster is $N_k = \sum_{i=1}^{N} \mathcal{I}(\pi_i = k)$, where $\mathcal{I}(\cdot)$ is the indicator function. It holds that $\sum_{k=1}^{K} N_k = N$. To simplify the notation it is convenient to allow for empty clusters, so that $K \lesseqgtr N$. We denote by μ_k the vector of GARCH parameters common to the assets included in the k-th cluster. By definition, it holds that $C_i = \mu_{\pi_i}$, $i = 1, 2, \ldots, N$. The set of VC-cDCC GARCH dynamic parameters is thus given by μ_i, $i = 1, 2, \ldots, K$. If $\pi_i \neq \pi_j$ for $i < j = 2, \ldots, N$, that is, no common GARCH dynamics, we get the cDCC model, which is therefore a special case of the VC-cDCC model. [3]

2.3 VC-cDCC Estimation

We assume that P and Q are known (typically $P = Q = 1$), and we arrange the unknown VC-cDCC parameters

$$(\tau, S, C, \phi, \pi),$$

where $\tau = [\tau_1, \tau_2, \ldots, \tau_N]'$ is the vector of the asset unconditional variances (see 4), $C = [C_1 : C_2 : \ldots : C_N]$ is the $(P + Q) \times N$ matrix of the GARCH dynamic parameters,[4] where : is the horizontal matrix concatenation operator, and ϕ is the vector stacking the distinct elements of the correlation dynamic parameter matrices, (A, B). The notation we adopt for the representation of the entire set of model parameters is redundant. In fact, the parameter matrix C contains replicated columns if the selected assets have a clustering structure. Nevertheless, to simplify the notation and the description of the estimation step, we maintain the matrix C, while we stress that the VC-cDCC GARCH parameters are those in the vectors μ_i, $i = 1, 2, \ldots, K$. We also note that in the limiting case of the cDCC model, all columns of C are different given the assets do not show evidence of a clustering structure.

Unless N is very small, the joint QML estimation of the VC-cDCC variance and correlation parameters is infeasible. The problem is in fact even more complex, given that the partition vector must also be estimated. As a feasible though no longer efficient estimation strategy, we consider a

[3]As we will show in the next sections, given a sample of assets, we estimate the number of clusters. Therefore, the cDCC model is one of the possible outcome in which all clusters are singleton.

[4]If the GARCH orders P and Q differ across the N assets, the missing elements are replaced by zeros in the C matrix.

two-step approach, called VC-cDCC estimator for short. The first step focuses on the estimation of the partition vector, denoted as $\hat{\pi} = [\hat{\pi}_1, \hat{\pi}_2, \dots, \hat{\pi}_N]'$, on the basis of the univariate GARCH estimates of the parameters in C. The second step focuses on the estimation of the cDCC model of Aielli (2006) subject to

$$C_i = C_j \text{ if } \hat{\pi}_i = \hat{\pi}_j, \tag{9}$$

which are the constraints induced by the clustering structure estimated in the first step. The above equation is thus the estimated counterpart of the variance clustering constraint of Eq. 8. Thanks to variance clustering, even for large N, we are able to jointly estimate the whole set of model dynamic parameters (variance and correlation dynamic parameters). This is a relevant advantage against the traditional cDCC estimator for large systems, where the variance process and the correlation process are separately estimated. In the following sections, we address the estimation of the partition vector and of the other model parameters, supporting the proposed approach with a simulation study.

3 Estimating the Partition Vector

The VC-cDCC model assumes that a set of assets is clustered with respect to the dynamic evolution of their GARCH variances, as summarized by the univariate GARCH parameters (intercept excluded). Thus, we face the problem of grouping the assets starting from the sample of the univariate GARCH estimates. One possible objection to our final purpose is the lack of clustering patterns across estimated GARCH parameters. For instance, the first panel of Figure 1 reports the scatter plot of the univariate QML estimates of GARCH(1,1) parameters on 100 stock returns series. The graph does not show any clear cluster. The second panel of Figure 1 is very similar to the first one, but derives from 100 simulated series clustered into four groups. This result is due to the estimation error of the GARCH parameters, and thus on their dispersion around the true data generating process values. Therefore, the estimation error could be responsible for the *apparent* absence of clusters within a set of GARCH(1,1) estimated parameters. Furthermore, we know that in large samples, the estimation error of the univariate QML GARCH estimator is approximately Gaussian. This fact provides a strong motivation for constructing an *ad hoc* algorithm for the clustering of GARCH processes based on Gaussian mixture (GM) models. The algorithm we propose below, called the GARCH-GM clustering algorithm, provides an estimate of the groups and thus of the partition vector π. The clustering algorithm we propose is also robust to the limiting case of absence of clusters (which leads to the cDCC model). In that case, the clustering algorithm will return a partition vector identifying a set of singleton clusters.

3.1 The GARCH-GM Clustering Algorithm

Denote as $(\tilde{\tau}_i, \tilde{C}_i)$ the univariate QMLE of the GARCH parameter of the i-th asset. We then model the entire set of the $\tilde{C}_i$'s as *iid* with GM pdf

$$\sum_{k=1}^{K} \omega_k \times f(\tilde{C}_i; \mu_k, \tilde{\Sigma}(\mu_k)/T), \; i = 1, 2, \ldots, N, \tag{10}$$

where T is the length of the series, $f(\cdot; \mu, \Xi)$ is the Gaussian pdf with mean μ and covariance matrix Ξ, and ω_k, $k = 1, 2, ...K$ are the mixture weights satisfying $\sum_{k=1}^{K} \omega_k = 1$. Denoting as

$$\{(\tilde{\omega}_k, \tilde{\mu}_k), k = 1, \ldots, \hat{K}\} \tag{11}$$

an estimate of the parameters in (10), the partition vector is estimated through the GM clustering outputs by setting

$$\hat{\pi}_i = \text{argmax}_{k=1,2,\ldots,\hat{K}} \; f(\tilde{C}_i; \tilde{\mu}_k, \tilde{\Sigma}(\tilde{\mu}_k)/T). \tag{12}$$

Here, $\hat{\pi}_i$, is the index of the Gaussian component with the largest contribution to the GM likelihood when evaluated at $\tilde{C}_i$ (see, among others, Xu and Wunsch, 2009, and therein cited references). Note that the component covariance matrix, $\tilde{\Sigma}(\mu_k)$, is modeled as a function of the mean of the component. The parameters in (12) are thus the means, the weights and the number of the components (there are no component covariance parameters to estimate). We stress that the number of components, K, is assumed unknown.

We motivate now the choice of (10) as a model for the $\tilde{C}_i$'s. Under general conditions (see Elie and Jeantheau 1995) the estimation error,[5] $\tilde{C}_i - C_i^0$, satisfies

$$\sqrt{T}(\tilde{C}_i - C_i^0) \overset{A}{\sim} N(0, \Sigma(C_i^0)), \tag{13}$$

where we assume that the asymptotic covariance matrix, $\Sigma(C_i^0)$, depends on C_i^0 only (this assumption is satisfied, for instance, if C_i and τ_i are orthogonal). Suppose, then, that: i) the return processes are independent, ii) the assets are drawn at random from the considered market, and iii) $Pr(C_i^0 = \mu_k^0) = \omega_k^0$, $k = 1, 2, \ldots, K$ (this is the true distribution of the GARCH dynamic parameters in the considered market), and iv) Eq. 13 holds. Then, if $\tilde{\Sigma}(\cdot)$ is a good approximation of $\Sigma(\cdot)$, it follows that, for large T, the $\tilde{C}_i$'s are approximately *iid* with GM probability density function given by (10). Assumption iii), which implies a discrete GARCH dynamic distribution within

[5] A zero superscript denotes true values.

the market, is not restrictive in that for sufficiently large K, we can approximate any continuous distribution on the $(P+Q)$-simplex of $R^{(P+Q+1)}$.[6] However, assumptions i-ii (also adopted in the clustering approach of Bauwens and Rombouts, 2007) are typically violated in financial applications: assets are normally correlated and, for instance, portfolios are not randomly created. In the following section, we present a simulation study showing that the performances of the GARCH-GM clustering algorithm turn out to be largely unaffected by violations of assumptions i and ii.

Given the model construction, to determine the optimal values (11) a BIC-based selection procedure is a logical choice. Compute the MLEs of (10) for a set of possible values of K and denote them as

$$\{(\tilde{\omega}_k^K, \tilde{\mu}_k^K), k = 1, \ldots, K\}. \tag{14}$$

Then we estimate K as

$$\hat{K} = \min\{K = 1, 2, \ldots : BIC_K < BIC_{K+1}\}, \tag{15}$$

where

$$BIC_K = -2\sum_{i=1}^{N} \log \sum_{k=1}^{K} \tilde{\omega}_k^K \times f(\tilde{C}_i; \tilde{\mu}_k^K, \tilde{\Sigma}(\mu_k^K)/T) \; + \; \{(K-1) + K(P+Q)\} \times \log N. \tag{16}$$

Finally, given the estimated $\hat{K}$ we set

$$(\tilde{\omega}_k, \tilde{\mu}_k) = (\tilde{\omega}_k^{\hat{K}}, \tilde{\mu}_k^{\hat{K}}), \;\; k = 1, 2, \ldots, \hat{K}.$$

The definition adopted for $\hat{K}$ provides a stopping rule that activates when $BIC_{\hat{K}+1}$ is reached. If BIC_K is a convex function of K, then $\hat{K}$ is a global minimizer of BIC_K. Furthermore, for small $\hat{K}$ (as it is typically the case; see section 5.3) this stopping rule also induces a computational advantage since it strongly reduces the estimation time. If BIC_K is infeasible (for example, for $K > N$, when the GM log-likelihood is ill-conditioned), we set $BIC_K = \infty$. Recalling that there are no parameters to take into account for the component covariance matrices, the number of parameters in the BIC formula, $(K-1)+K(P+Q)$, is equal to the number of free mixture weights plus the number of parameters in the component means.

In order to complete the description of the clustering algorithm, we have to define $\tilde{\Sigma}(\cdot)$. An appropriate specification of $\tilde{\Sigma}(\cdot)$ should be feasible and possess desirable analytical and/or statistical properties. As a possible solution, we propose to set

$$\tilde{\Sigma}(x) = \sum_{i=1}^{N} \psi_i(x) \times \tilde{\Sigma}_i, \qquad \psi_i(x) = ||x - \tilde{C}_i||^{-1} \Big/ \sum_{i=1}^{N} ||x - \tilde{C}_i||^{-1}, \tag{17}$$

[6] Recall that the $P+Q$ GARCH dynamic parameters are non-negative and that their sum must be less than one.

where $||\cdot||$ is the Euclidean norm and $\tilde{\Sigma}_i$ is the QML sandwich estimator of the asymptotic covariance matrix of $\tilde{C}_i$. We stress that, by this specification of $\tilde{\Sigma}(\cdot)$, all inputs we need in order to run our clustering algorithm are the estimated univariate GARCH dynamic parameters and their corresponding estimated covariance matrices. Noting that $\psi_i(x) \geq 0$ and that $\sum_{i=1}^{N} \psi_i(x) = 1$, it follows that $\tilde{\Sigma}(x)$ is a weighted mean of $\tilde{\Sigma}_1, \ldots, \tilde{\Sigma}_N$. By inspection of the formula of $\psi_i(x)$, it is easily proved that, under consistency of $(\tilde{C}_i, \tilde{\Sigma}_i)$, it holds

$$\operatorname{plim}_{T \longrightarrow \infty}, \tilde{\Sigma}(\mu^0_{\pi_i}) = \Sigma(\mu^0_{\pi_i}), \tag{18}$$

where $i = 1, 2, \ldots, N$, that is, $\tilde{\Sigma}(\mu^0_{\pi_i})$ is consistent, as $T \longrightarrow \infty$, if evaluated at the true value of the component mean. Hence, for sufficiently large T, locally, in a neighborhood of the true component mean, $\tilde{\Sigma}(\cdot)$ is a good approximation of $\Sigma(\cdot)$. By construction, $\tilde{\Sigma}(x)$ is strictly positive definite uniformly on the $(P+Q)$-simplex of $R^{(P+Q+1)}$. Thus, compared with other traditional specifications of the component covariance matrix such as the heteroskedastic GM clustering (see below), the GM log-likelihood keeps well-conditioned irrespective of critical quantities; the minimum cluster size, for example. As an important consequence of this property, our clustering algorithm remains feasible even in the case of no common GARCH dynamics, $\hat{K} \geq N$. Furthermore, we notice that, for increasing T, the dispersion of the component decreases, which means that the accuracy of the estimated partition vector increases with T. It is worth noting that, for such a property to hold, the number of assets does not need to increase (N is always kept fixed in the discussion above).

3.2 Simulation Study

In this section, we compare the performances of the GARCH-GM clustering algorithm with those of other simpler GM clustering techniques, namely, heteroskedastic (HE) GM clustering and homoskedastic (HO) GM clustering. HE-GM clustering is computed replacing $\Sigma(\mu_k)$ in model (10) with a free, positive semi-definite, component-specific parameter matrix Σ_k, $k = 1, 2, \ldots, K$. The number of parameters entering the BIC formula in this case is $(K-1)+K(P+Q)+K(P+Q)(P+Q+1)/2$. In contrast, HO-GM clustering is computed replacing $\Sigma(\mu_k)$ in (10) with a free, positive semi-definite parameter matrix Σ, common to all components. The number of distinct parameters entering the BIC formula in this case is $(K-1)+K(P+Q)+(P+Q)(P+Q+1)/2$. To maximize the univariate GARCH quasi-log-likelihood, we adopt a grid-based choice for the starting values. This procedure reduces the risk of choosing as optimal parameters those associated with a local maxima; see Paolella (2010) for a discussion on the local maxima in GARCH QML estimation.

3.2.1 Data Generating Processes

For the simulation experiments, we consider two data generating processes, one for large datasets, where $N = 100$, and another one for small datasets, where $N = 10$. The variance processes are assumed GARCH(1,1) with parameters set as in Table 1. These parameter settings are coherent with the estimation output provided by the real datasets employed in this paper (see section 5.1). To introduce dependence across the simulated assets, we assume constant conditional correlations, which we set as

$$s_{ij} = s = 0, .3, .6, .9, \ \ i < j = 2, \ldots, N.$$

Forcing s to vary in a range of admissible values allows us to check whether the GARCH-GM clustering performances are affected by violations of the assumption of independence across the return processes. We then generate the variance standardized returns as $\varepsilon_t = R_t^{1/2} z_t$, where $R_t^{1/2}$ is Cholesky's square root of R_t, and where the elements of z_t are *iid* standardized Student's t-distributed with $\nu = 9$ degrees of freedom. We preferred the Student density to be consistent with the empirical evidences of leptokurtosis in financial returns series. For each parameter constellation, we generate 500 independent samples. Furthermore, to limit the computation time, we run the BIC-selection procedure only for $K = 1, 2, 3, 5, 8, 13$ on the large datasets, and only for $K = 1, 2, 3, 5, 8$ on the small datasets.

3.2.2 Performance Criteria

We evaluate the performances of the clustering algorithm by comparing the estimated number of clusters, $\hat{K}$, with the true number of clusters, K^0. Furthermore, we consider a measure of incorrect partition, which we define as

$$\hat{\Psi} = \frac{1}{N(N+1)/2} \sum_{i<j=2,\ldots,N} |p_{ij}^0 - \hat{p}_{ij}|,$$

where

$$\begin{cases} p_{ij}^0 & = & \mathcal{I}(\pi_i^0 = \pi_j^0); \\ \hat{p}_{ij} & = & \mathcal{I}(\hat{\pi}_i = \hat{\pi}_j \,). \end{cases} \tag{19}$$

A smaller $\hat{\Psi}$ is associated to a better fit of $\hat{\pi}$ to π^0. It holds $0 \leq \hat{\Psi} \leq 1$, where $\hat{\Psi} = 0$ if and only if $\hat{\pi} = \pi^0$. The actual maximum of $\hat{\Psi}$ is less than one and it depends on π^0. We note that $\hat{\Psi}$ does not require that the compared partitions need have the same number of clusters. See Xu and Wunsch (2009) for other approaches and methods to evaluate the performances of clustering algorithms.

3.2.3 Results

An illustrative plot relative to the considered clustering methods is reported in Figure 2, while additional graphs are included in the Appendix. We start from the the GARCH-GM clustering (see Figure 2). As expected, for small sample sizes ($T = 250$), the dispersion of the univariate QML estimators is too large for the asset GARCH dynamics to be correctly partitioned. The estimated number of clusters is incorrect (the smallest BIC is for $K = 3$, whereas the true number of clusters is $K^0 = 2$). On the other hand, in the cases of $T = 1250$ and $T = 2500$, the number of clusters is correctly detected, and the asset GARCH dynamics are accurately partitioned. By comparing the shape and the dispersion of the component contour plots with the related clusters, the choice of $\tilde{\Sigma}(\tilde{\mu}_k)$ as component covariance matrix estimator seems appropriate.

Moving to HE-GM clustering, we note that the component covariance matrix is correctly specified but, unfortunately, for this clustering technique to be feasible, it is required that $\hat{K}$ is not too large with respect to N, and that the cluster sizes are not too small.[7] Figure A.1 illustrates the behavior of the HE-GM clustering relying on the same simulated dataset used with the GARCH-GM clustering. In spite of the correct specification of the component covariance matrix, some odd behaviors can arise. In fact, there are positively correlated clusters whose shape is opposite the stochastic properties of the GARCH(1,1) QMLE. However, for $T = 2500$ the number of clusters is correctly detected (the smallest BIC is for $K = 2 = K^0$), and the asset GARCH dynamics are accurately partitioned.

Finally, in the HO-GM clustering, we observe that, thanks to homoskedasticity, this procedure remains feasible even for small cluster sizes. However, the misspecification of the component covariance matrix can be substantial (see Figure A.2), as shown by the presence of positively correlated clusters, and by the fact that many observations fall outside the 90% component contour plots. Even for large T the number of clusters is overestimated.

By analyzing in greater detail the simulation results of the large dataset experiment (see Figure 3), we note that GARCH-GM clustering and HE-GM clustering have similar performances. However, there is a slight preference for the GARCH-GM clustering in terms of percentage of incorrect allocations. In contrast, the HE-GM clustering is slightly preferred in terms of estimated number of clusters. Note that, as long as the asset cross-correlation increases, the performances of both clustering algorithms improve. For $s = 0.9$, the GARCH-GM clustering globally outperforms HE-GM clustering in terms of both incorrect partition and estimated number of clusters. The

[7] As a consequence, the HE-GM clustering is not robust to the absence of clusters.

performances of HO-GM clustering in terms of incorrect allocations are satisfactory, and they are almost unaffected by changes in the asset cross-correlation. In terms of the estimated number of clusters, however, HO-GM clustering exhibits a marked tendency to overestimate the true value.

Moving to the small dataset experiment (see Figure 4), GARCH-GM clustering strongly outperforms HO-GM clustering in terms of both incorrect allocations and estimated number of clusters. The HO-GM clustering confirms the previous results, that is, the overestimation of the true number of clusters. Because of the large number of parameters to estimate with respect to the number of observations, HE-GM clustering is infeasible in this experiment.

In summary, for reasons of flexibility and the performances in both the simulation experiments, GARCH-GM clustering seems to be globally the best method among the three considered clustering techniques. In the case of large datasets with approximately balanced clusters, HE-GM clustering can be a valid alternative to the GARCH-GM clustering. However, we stress that *a-priori* the number of clusters and their sizes are not known, an element favouring the use of the GARCH-GM clustering method.

3.3 Consistency

As a further issue, we are interested in evaluating the consistency of the estimated partition vector for $T \longrightarrow \infty$ given a set of selected assets. We do not provide a rigorous proof of consistency, but confine ourself to noting that, under consistency of the univariate GARCH QMLEs (see Elie and Jeantheau, 1995), for increasing T, the $\tilde{C}_i$s concentrate around their true values with increasing probability. Therefore, for a fixed set of assets (given a choice of N) and for sufficiently large T, any "well-constructed" clustering algorithm should be capable of correctly detecting the true underlying GARCH dynamic partition. Suppose as an ideal example that there are exactly N assets; those assets have Ndistinct GARCH dynamics that spread out according to a Gaussian scatter-plot. For large T, the univariate GARCH estimators are close to the true values and, then they too spread out according to a Gaussian scatter-plot. In this case, both HO-GM clustering and HE-GM clustering outcomes are driven by the Gaussianity of the scatter-plot of the univariate GARCH estimates. Particularly, they both select the *incorrect* one-cluster partition as the best partition. In contrast, since the dispersion of the GARCH-GM component is a decreasing function of T, for large T the GARCH-GM clustering is capable of correctly detecting the absence of clusters. An example is given in Figure A.3, where for $T = 1250$ and $T = 2500$ the absence of clusters is correctly detected, illustrates this property.

4 Estimating the Constrained cDCC Model

The second step of the VC-cDCC estimator focuses on the estimation of the cDCC model subject to the equality constraints in Eq. 9. Even under such equality constraints, unless N is very small, the dimension of the cDCC parameter space is too large for the joint QML estimator always to be feasible. This limitation depends mainly on the need to estimate the elements in the parameter matrix S (that enter in the intercept of the dynamic correlations). In fact, S contains $N(N-1)/2$ parameters, and is thus of order $O(N^2)$. On the other hand, the remaining model parameters include N unconditional variances, $2K$ dynamic GARCH parameters, and 2 parameters in the case of Scalar cDCC parametrization. Following Aielli (2006) we restore feasibility by resorting to a generalized profile QML estimator.

4.1 Generalized Profile Dynamic Parameter Estimation

In cDCC estimation, we treat the dynamic parameters, (C, ϕ), as parameters of interest and the intercept parameters, (τ, S), as nuisance parameters. Note that our purpose is to jointly estimate both the GARCH and correlation dynamic parameters. Denote the cDCC quasi-log-likelihood as $L(\tau, S, C, \phi)$, and suppose the availability of an estimator of (τ, S) conditionally on (C, ϕ), denoted as $(\hat{\tau}_{(C,\phi)}, \hat{S}_{(C,\phi)})$. The function, $L(\hat{\tau}_{(C,\phi)}, \hat{S}_{(C,\phi)}, C, \phi)$, called a generalized profile quasi-log-likelihood (Severini 1998), is a function of (C, ϕ) only. The estimator of (C, ϕ) is

$$(\hat{C}, \hat{\phi}) = \text{argmax}_{(C,\phi)}\, L(\hat{\tau}_{(C,\phi)}, \hat{S}_{(C,\phi)}, C, \phi) \tag{20}$$

subject to $C_i = C_j$ if $\hat{\pi}_i = \hat{\pi}_j$, $i, j = 1, 2, \ldots, N$, $i \neq j$. The (unconditional) estimator of (τ, S), is

$$(\hat{\tau}, \hat{S}) = \big(\hat{\tau}_{(\hat{C},\hat{\phi})}, \hat{S}_{(\hat{C},\hat{\phi})}\big), \tag{21}$$

which is the same as the value of the conditional estimator of (τ, S) at the end of the generalized profile quasi-log-likelihood maximization.

The feasibility of $(\hat{\tau}, \hat{S}, \hat{C}, \hat{\phi})$ depends on the number of parameters of the generalized profile quasi-log-likelihood, which is $(P+Q)\hat{K}+\dim(\phi)$, and on the feasibility of the conditional estimator. Note that, for large N, as in our focus, $\hat{K}$ turns out to be typically small (e.g., $\hat{K} \leq 5$ with $N = 100$ — see section 5). Hence, for suitably restricted ϕ (e.g., under scalar correlation dynamics) the computation of $(\hat{C}, \hat{\phi})$ does not pose any particular problem. A standard choice for the conditional estimator would be to define $(\hat{\tau}_{(C,\phi)}, \hat{S}_{(C,\phi)})$ as the QML estimator of (τ, S) conditionally on (C, ϕ), namely,

$$(\hat{\tau}_{(C,\phi)}, \hat{S}_{(C,\phi)}) = \text{argmax}_{(\tau,S)}\, L(\tau, S, C, \phi). \tag{22}$$

In this case, the objective function in Eq. 20 is the traditional profile quasi-log-likelihood. Unfortunately, the computation of such a conditional estimator requires the solution of a highly non-linear $O(N^2)$ maximization problem, and, then it is infeasible for large N. As a feasible conditional estimator, we suggest the following specification, which is constructed relying on the property of the cDCC model given in Eq. 7. For fixed (C, ϕ),

1) set $\hat{\tau}_{(C,\phi)} = \tilde{\tau} = [\tilde{\tau}_1, \ldots, \tilde{\tau}_N]'$;

2) set $\hat{S}_{(C,\phi)} = T^{-1} \sum_{t=1}^{T} \{\tilde{Q}_t^* \tilde{\varepsilon}_t\}\{\tilde{Q}_t^* \tilde{\varepsilon}_t\}'$, where

    (a) $\tilde{\varepsilon}_t^* = [\tilde{\varepsilon}_{1,t}^*, \ldots, \tilde{\varepsilon}_{N,t}^*]'$;

    (b) $\tilde{\varepsilon}_{1,t}^* = y_{i,t}/\sqrt{\tilde{h}_{1,t}}$;

    (c) $\tilde{h}_{i,t} = (1 - \sum_{p=1}^{P} a_{ip} - \sum_{q=1}^{Q} b_{i,q})\, \tilde{\tau}_i + \sum_{p=1}^{P} a_{i,p}\, y_{i,t-p}^2 + \sum_{q=1}^{Q} b_{i,q}\, \tilde{h}_{i,t-q}$;

    (d) $\tilde{Q}_t^* = \mathrm{diag}(\sqrt{\tilde{q}_{1,t}}, \ldots, \sqrt{\tilde{q}_{N,t}})$;

    (e) $\tilde{q}_{i,t} = (1 - \alpha_{ii} - \beta_{ii}) + \alpha_{ii}\, \tilde{\varepsilon}_{it-1}^2 \tilde{q}_{i,t-1} + \beta_{ii}\, \tilde{q}_{i,t-1}$.

The estimator $\hat{\tau}_{(C,\phi)} = \tilde{\tau}$ in 1) is the vector of the univariate QML intercept parameter estimators computed only once at the beginning of the GARCH-GM clustering algorithm. Since $\tilde{\tau}$ does not depend on (C, ϕ), it does not need to be recomputed at each evaluation of the generalized profile quasi-log-likelihood. The estimator $\hat{S}_{(C,\phi)}$ is defined as the sample counterpart of S evaluated at (C, ϕ) (see Eq. 7). Since $\hat{S}_{(C,\phi)}$ depends on (C, ϕ), it must be recomputed at each evaluation of the generalized profile quasi-log-likelihood. Note, however, that the recursions required by the evaluation of $\hat{S}_{(C,\phi)}$ are not computationally complex. Furthermore, if $Q_t^* \varepsilon_t$ is second moment ergodic (see Aielli, 2011), for known τ^0 the estimator $\hat{S}_{(C,\phi)}$ satisfies

$$\operatorname{plim} \hat{S}_{(C^0,\phi^0)} = S^0. \tag{23}$$

This property is a direct consequence of Eq. 7, and of the fact that, by definition, $\hat{S}_{(C^0,\phi^0)}$ is the sample covariance matrix of the true unobserved vector $Q_t^* \varepsilon_t$. Building on this property, a heuristic proof of consistency of the VC-cDCC estimator will be provided in section 4.3.

We now discuss three possible variants of the estimator just introduced.

4.1.1 Feasibility Irrespective of $\hat{K}$

A VC-cDCC estimator that is feasible irrespective of $\hat{K}$ can be obtained treating (τ, C, S) as a nuisance parameter and ϕ as a parameter of interest. Denoted as $(\hat{\tau}_\phi, \hat{S}_\phi, \hat{C}_\phi)$ the conditional estimator of the nuisance parameter, where $\hat{C}_\phi$ is subject to (9), the estimator of ϕ is computed as

$$\hat{\phi} = \text{argmax}_\phi \, L(\hat{\tau}_\phi, \hat{S}_\phi, \hat{C}_\phi, \phi), \tag{24}$$

and the unconditional estimator of (τ, S, C) is computed as

$$(\hat{\tau}, \hat{S}, \hat{C}) = (\hat{\tau}_{\hat{\phi}}, \hat{S}_{\hat{\phi}}, \hat{C}_{\hat{\phi}}). \tag{25}$$

A feasible specification of $(\hat{\tau}_\phi, \hat{S}_\phi, \hat{C}_\phi)$ is the following one, where $L_i(\tau_i, C_i)$ denotes the univariate quasi-log-likelihood of y_{it}. For fixed ϕ,

i) set $\hat{\tau}_\phi = \tilde{\tau}$;

ii) set $\hat{C}_\phi = \hat{C} = [\hat{\mu}_{\hat{\pi}_1} : \hat{\mu}_{\hat{\pi}_2} : \ldots : \hat{\mu}_{\hat{\pi}_N}]$, where

$$\hat{\mu}_k = \text{argmax}_\mu \sum_{i=1}^{N} \mathcal{I}(\hat{\pi}_i = k) \times L_i(\tilde{\tau}_i, \mu), \; k = 1, \ldots, \hat{K};$$

iii) set $\hat{S}_\phi = \hat{S}_{(\hat{C},\phi)}$, where $\hat{S}_{(C,\phi)}$ is defined as in point 2 of section 4.1.

The main difference between the estimator defined by Eqs. 20-21, points 1-2 (hereafter, *joint* VC-cDCC estimator) and the estimator defined by equations (24-25), points i-iii (hereafter, *sequential* VC-cDCC estimator), is that, with the joint estimator, all dynamic parameters of the model are estimated jointly, whereas, with the sequential estimator, the dynamic variance parameters and the dynamic correlation parameters are estimated in subsequent steps. The calculations required by the sequential estimator are much faster than those required by the joint estimator. In particular, the computation of $\hat{C}_\phi = \hat{C}$ reduces to $\hat{K}$ small size maximizations,[8] where the k-th objective function is the sum of the univariate quasi-log-likelihoods of the assets belonging to the k-th cluster, also called *univariate composite* likelihood (see Pakel et al., 2011). In the case of scalar correlation dynamics, the generalized profile quasi-log-likelihood in Eq. 24 turns out to be a function of two variables only, irrespective of N. For $\hat{K} \geq N$ (that is, possibly no common GARCH dynamics) the sequential estimator coincides with the traditional large system cDCC estimator (Aielli, 2006). According to the notation adopted in this paper, the cDCC estimator is computed replacing step ii) above with

ii') set $\hat{C}_\phi = \hat{C} = [\tilde{C}_1 : \tilde{C}_2 : \ldots : \tilde{C}_N]$.

Finally, we note that, from a computational point of view, both the joint and the sequential estimators benefit from the availability of a good starting value for the parameters in C, that is, the set of the component means provided as an output of the GARCH-GM clustering.

[8]Each maximization determines $P + Q$ parameters.

4.1.2 Bivariate Composite Likelihood

As shown by Engle et al. (2009), especially for large N, profiling the cDCC quasi-likelihood, as in the joint estimator and in the sequential estimator described above, can result in a large bias of the estimated correlation innovation parameters. This problem can be strongly alleviated by profiling the so-called *bivariate composite* cDCC quasi-log-likelihood. We thus replace $L(\tau, S, C, \phi)$ in Eq. 20 and Eq. 24 with

$$\sum_{i=2}^{N} L_{i,i-1}(\tau, S, C, \phi), \tag{26}$$

where $L_{i,j}(\tau, S, C, \phi)$ is the quasi-log-likelihood of the cDCC submodel of $(y_{i,t}, y_{j,t})$. The resulting approach is referred to as the bivariate composite likelihood approach (see Engle et al., 2009, for a general discussion). With respect to the calculations involved by the full cDCC quasi-log-likelihood, the adoption of the bivariate composite quasi-log-likelihood results in a dramatic reduction of CPU time, especially for large N.

4.1.3 Variance Targeting.

Recalling that τ_i is the unconditional variance of y_{it} (see section 2.1), $\tilde{\tau}$ in points 1 and i can be replaced by $\hat{\tau} = [\hat{\tau}_1, \ldots, \hat{\tau}_N]'$, where

$$\hat{\tau}_i = T^{-1} \sum_{t=1}^{T} y_{i,t}^2. \tag{27}$$

Such an estimation device is known as variance targeting (Engle and Mezrich, 1996). Variance targeting is proved to have good empirical performance, especially under misspecifications (see Francq et al., 2009). Hereafter, if $\hat{\tau}$ is used in place of $\tilde{\tau}$, then the VC-cDCC estimator is called *targeted* VC-cDCC estimator.

4.2 Simulation Study

In this section, using Monte Carlo methods, we compare the VC-cDCC estimators described above and the traditional cDCC estimator. The considered estimators are computed in the related bivariate composite versions. In summary, we find that the joint VC-cDCC estimator and the sequential VC-cDCC estimator both perform similarly in practice and better than the cDCC estimator. The targeted VC-cDCC estimators are recommendable with respect to the corresponding non-targeted versions, especially when dealing with large systems. The fact that the joint estimator and the sequential estimator are empirically equivalent means that there is no sensible loss of efficiency due to estimating the model's dynamic parameters in two steps. We stress that the sequential estimator is extremely rapid to compute and is feasible irrespective of $\hat{K}$.

We consider two simulation experiments, one for large N and another for small N, both with GARCH parameters and t-distributed returns as in section 3.2.1. A scalar dynamic governs the correlation process with parameters

$$\alpha = .02;\ \beta = .97;\ s_{ij} = s\ = .3\ \ i < j = 2, \ldots, N.$$

These correlation settings are coherent with typical cDCC estimation outputs. For each experiment, we generate $M = 500$ series of length $T = 1250$.

To measure the performances of the estimated variance process, we adopt the following mean absolute errors

$$VARMAE = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{h}_{i,T+1} - h_{i,T+1} \right| \tag{28}$$

and

$$CORRMAE = \frac{1}{N(N-1)/2} \sum_{i<j=2,\ldots,N} \left| \hat{\rho}_{ij,T+1} - \rho_{ij,T+1} \right|, \tag{29}$$

where $\hat{h}_{i,T+1}$ and $\hat{\rho}_{ij,T+1}$ are the out-of-sample estimates of, $h_{i,T+1}$ and $\rho_{ij,T+1}$, respectively, based on $y_1, y_2, \ldots, y_T$. We also compare the density plots of the parameter estimators.

4.2.1 Results

Figure 5 refers to the large dataset experiment. All plots show that the joint VC-cDCC estimator and the corresponding sequential VC-cDCC estimators provide the same empirical performances in practice. On the contrary, the performances of the targeted VC-cDCC estimators and of the corresponding non-targeted versions can be different, with a preference for the targeted estimator. Consider the variance results in plot (1): the VC-cDCC out-of-sample predictions, both targeted and non-targeted, perform better than the cDCC variance prediction (the $VARMAE$ density of the VC-cDCC estimators is more shifted towards zero than the corresponding cDCC density). As expected, thanks to variance clustering, the VC-cDCC parameter estimators are more concentrated around the true values than the corresponding cDCC estimators (see plots [3], [5] and [7] in Figure 5). Because of the estimation error in the partition vector, the behavior of the VC-cDCC estimator can be less regular than the one of the cDCC estimator, as shown by the bimodality of the VC-cDCC estimators of b_i (see plot [7]). The targeted estimators are generally well centered around the true values, whereas the non-targeted versions can exhibit rather large relative biases (see plots [3] and [7]). Focusing now on the correlation results, there is no apparent increase of prediction efficiency due to the variance clustering (see plot [2]). As for the correlation parameters, the targeted VC-cDCC density plots and the corresponding cDCC density plots are well centered on the true values and they are almost identical in practice (see plots [4], [6] and [8]). The non-targeted VC-cDCC estimators instead exhibit a rather large relative bias in the estimation of α (see plot [6]).

The simulation results for small N are reported in Figure A.4 in the Appendix. They confirm the fact that there is little difference between the joint VC-cDCC estimation and the sequential VC-cDCC estimation. The best variance predictions here are provided not only by the non-targeted VC-cDCC estimators, but also the targeted VC-cDCC estimators outperform the cDCC estimator (see plot [1]). As expected, in the case of the second cluster, which includes one asset only, the estimation error in the partition vector makes the VC-cDCC variance parameter estimators performances worse than the corresponding cDCC estimator (see the right quadrants in plots [5] and [7]). As noted in the case of the large dataset, the correlation performances appear to be unaffected by allowing for variance clustering (see plots [2], [4], [6] and [8]).

4.3 Consistency

A rigorous proof of consistency of the VC-cDCC estimator is quite difficult to provide. Here we confine ourself to sketch a heuristic proof that holds under high level assumptions. Consider the joint VC-cDCC estimator and suppose, for simplicity, that (π^0, τ^0) is known. The cDCC quasi-log-likelihood is then denoted as $L(S, C, \phi)$. Suppose that (i) the limits in probability, $\operatorname{plim} \hat{S}_{(C,\phi)}$ and $\operatorname{plim} T^{-1} L(S, C, \phi)$, are finite in a neighborhood of (C^0, ϕ^0) and (S^0, C^0, ϕ^0), respectively, and (ii) $\operatorname{plim}\ T^{-1} L(S, C, \phi)$ has a unique maximum at (S^0, C^0, ϕ^0) (the latter is a common assumption in QML settings, which holds, e.g., under the conditions for QML consistency in Bollerslev and Wooldridge, 1992). Then

$$\begin{aligned} \operatorname{plim} T^{-1} L(\hat{S}_{(C,\phi)}, C, \phi) &= \operatorname{plim} T^{-1} L(\operatorname{plim} \hat{S}_{(C,\phi)}, C, \phi) \\ &\leq \operatorname{plim} T^{-1} L(S^0, C^0, \phi^0), \end{aligned} \tag{30}$$

where the equality follows from Wooldridge (1994, Lemma A.1). Now, recalling that $\operatorname{plim} \hat{S}_{(C^0,\phi^0)} = S^0$ (see Eq. 23), it follows that the rescaled generalized profile quasi-log-likelihood, $\operatorname{plim} T^{-1} L(\hat{S}_{(C,\phi)}, C, \phi)$, has a unique maximum in (C^0, ϕ^0). This proves the local consistency of $(\hat{C}, \hat{\phi}) = \operatorname{argmax}_{(C,\phi)} L(\hat{S}_{(C,\phi)}, C, \phi)$ provided that point-wise convergence can be turned into uniform. As for consistency of $\hat{S}$, from Wooldridge (1994, Lemma A.1) and Eq. 23, it follows that, if $\operatorname{plim} \hat{S}_{(C,\phi)}$ is finite in a neighborhood of (C^0, ϕ^0) and $(\hat{C}, \hat{\phi})$ is consistent, then $\operatorname{plim} \hat{S} = \operatorname{plim} \hat{S}_{(\hat{C},\hat{\phi})} = \operatorname{plim} \hat{S}_{(C^0,\phi^0)} = S^0$. If (π^0, τ^0) is replaced by $(\hat{\pi}, \tilde{\tau})$ or $(\hat{\pi}, \hat{\tau})$, consistency of $(\hat{C}, \hat{\phi})$ holds under conditions for consistency of two-step M-estimators such as those given in Wooldridge (1994, Theorem 4.3). For consistency of $\hat{\pi}$, see section 3.3. Consistency of $\tilde{\tau}$ holds under Elie and Jeantheau (1995). If $\hat{\tau}$ is used in place of $\tilde{\tau}$, since $\hat{\tau}_i^2$ is the sample variance of $y_{i,t}$, consistency of $\hat{\tau}$ holds under second moment ergodicity of y_{it}, $i = 1, 2, \ldots, N$.

5 Applications to Real Data

In this section, we compare the empirical performances of the VC-cDCC estimator and those of the cDCC estimator. In summary, according to the considered performance criteria, we find that the targeted VC-cDCC estimators always perform either equal or better than the cDCC estimator. As anticipated by the simulation evidence, the difference between the joint VC-cDCC estimators and the corresponding sequential versions is negligible. In contrast to the simulation results, the estimation of the correlation process is markedly improved by allowing for variance clustering. This somewhat surprising outcome depends, in our opinion, on the rolling estimation strategy adopted. We believe that the targeted version is less sensible to large changes in the assets unconditional variances. Furthermore, within the simulations the unconditional variances remained fixed, while in the real data case, the unconditional variances might change over time, with a possible impact on the model ranking.

5.1 Datasets and Compared Estimators

We consider a large dataset of $N = 100$ randomly selected equities from the S&P 1500 Index. These equities belong either to the industrial or to the consumer good sectors, and their daily total returns are available from January 2001 to December 2007. Therefore, this dataset has an expectation on the number of clusters, which is equal to 2. The sample period thus includes 1770 daily returns. For the same period, we consider also a small dataset of equity indices: the nine S&P 500 SPDR's sector indices and the S&P 500 index itself (see Table A.10 in the Appendix for the detailed list of the considered assets). On the two datasets, we estimate both cDCC and VC-cDCC models. For the second model, we consider both the joint estimators of the number of clusters and of the other model parameters, denoted as $cDCC_{VC}$, as well as the estimators of model parameters conditionally on a *a priori* defined number of clusters denoted as $cDCC_K$. Estimating models with a fixed number of clusters allows a deeper understanding of the behavior of the VC-cDCC estimator. Another interesting question is whether the data driven partition of the

VC-cDCC estimator outperforms some relevant *a priori* partition criterion. To shed some light on this aspect, we also estimate on the large dataset a VC-cDCC model with *a priori* fixed partition vector denoted as $cDCC_{*}$. The *a priori* partition is given by the sector partition into industrial equities (53 assets) and consumer good equities (47 assets). For all of these estimators we also consider the related targeted version (see section 4.1.3) denoted by a superscript "Trg". Thus, for example, the targeted VC-cDCC estimator is denoted as $cDCC_{VC}^{Trg}$. We evaluate all estimators in their bivariate composite version (see section 4.1.2) assuming GARCH(1,1) variance processes and scalar correlation dynamics.

5.2 Performance Criteria

We compare the several model estimators by means of out-of-sample prediction criteria. In particular, we consider pairwise comparison tests of equal predictive ability (see Diebold and Mariano, 1995, and Patton and Sheppard, 2009, for applications in MGARCH models) and regression-based specification tests, such as the Engle-Colacito regression (Engle and Colacito, 2006), the dynamic quantile test (Engle and Manganelli, 2001), and the LM test of ARCH effects on portfolio returns. We consider as portfolio specifications the equally weighted portfolio, denoted as EW, the minimum variance portfolio with short selling, denoted as MV, and the minimum variance portfolio without short selling, denoted as MV*. The weight vector of the EW portfolio is $w_t^0 = w = \iota/N$. The weight vector of the MV portfolio is $w_t^0 = \{H_t^0\}^{-1}\iota/(\iota'\{H_t^0\}^{-1}\iota)$, and, then it is an explicit function of the conditional covariance matrix. Finally, the weight vector of the MV* portfolio is a non-closed function of the conditional covariance matrix, and it must be numerically computed. During the computation of the portfolio weights, H_t^0 is replaced by an estimate. Since the EW weights do not depend on H_t^0, they are free from estimation errors. In the following, we report a description of the adopted comparison criteria. The symbol $\hat{H}_t^{(m)}$ denotes the out-of-sample estimate of H_t^0, which is computed applying the m-th estimator on a rolling window of $\bar{t}$ de-meaned observed returns, $\hat{y}_{t-1}, \hat{y}_{t-2}, \ldots, \hat{y}_{t-\bar{t}}$, where $\hat{y}_{t-j} = z_{t-j} - \bar{z}_{t-1}$, and z_{t-j} is the vector of the observed return at time $t-j$ and $\bar{z}_{t-1} = \bar{t}^{-1}\sum_{j=1}^{\bar{t}} z_{t-j}$. We adopt a rolling window of length $\bar{t} = 1250$. The number of estimated forecasts for each estimator is then $\bar{T} = T - \bar{t} = 520$, roughly corresponding to two years of data.

5.2.1 Pairwise Comparison Tests of Equal Predictive Ability

Denote as $d_t^{(m)}$ the loss due to predicting H_t^0 with $\hat{H}_t^{(m)}$, and set

$$\bar{d}^{(m)} = \bar{T}^{-1} \sum_{t=\bar{t}+1}^{T} d_t^{(m)}.$$

The null hypothesis

$$\mathcal{H}^0 : E\left[\bar{d}^{(m)} - \bar{d}^{(n)}\right] = 0$$

is a null of equal predictive ability for the estimators $\hat{H}_t^{(m)}$ and $\hat{H}_t^{(n)}$. Under appropriate conditions (see Diebold and Mariano, 1995), a test statistic for $\mathcal{H}^0$ is

$$\frac{\sqrt{\bar{T}}(\bar{d}^{(m)} - \bar{d}^{(n)})}{\sqrt{VAR[\sqrt{\bar{T}}(\bar{d}^{(m)} - \bar{d}^{(n)})]}} \sim N(0,1),$$

where $VAR[\sqrt{\bar{T}}(\bar{d}_m - \bar{d}_n)]$ is a heteroskedasticity and autocorrelation consistent estimate of the variance of $\sqrt{\bar{T}}(\bar{d}_m - \bar{d}_n)$. Negative values of the test statistic provide evidence in favor of $\hat{H}_t^{(m)}$ against $\hat{H}_t^{(n)}$. As possible specifications of $d_t^{(m)}$ we consider score-based losses as in Amisano and Giacomini (2007), and MSE-based losses as in Diebold and Mariano (1995) and Patton and Sheppard (2009).

□ **Amisano-Giacomini losses.** Given our dynamic correlation setting, we decide to compute two separate score-based losses, one for the asset variance process and another for the asset correlation process, both computed under Gaussianity. The loss for the variance process is

$$d_t^{(m)} = VARSCORE_t^{(m)} = \sum_{i=1}^{N} \left(\log \hat{h}_{i,t}^{(m)} + \hat{y}_{i,t}^2 / \hat{h}_{i,t}^{(m)}\right), \tag{31}$$

where $\hat{h}_{it}^{(m)}$ is the out-of-sample estimate of h_{it}. The loss for the correlation process is

$$d_t^{(m)} = CORRSCORE_t^{(m)} = \log|\hat{R}_t^{(m)}| + \hat{\varepsilon}_t^{(m)\prime}\{\hat{R}_t^{(m)}\}^{-1}\hat{\varepsilon}_t^{(m)}, \tag{32}$$

where $\hat{\varepsilon}_t^{(m)} = \left[\hat{y}_{1,t}/\sqrt{\hat{h}_{1,t}^{(m)}}, \ldots, \hat{y}_{N,t}/\sqrt{\hat{h}_{N,t}^{(m)}}\right]'$. The loss for the correlation process depends on the estimated variances through $\hat{\varepsilon}_t^{(m)}$. We also consider a score-based loss for EW portfolio returns, defined as

$$d_t^{(m)} = EWSCORE_t^{(m)} = \log(w' \hat{H}_t^{(m)} w) + \left(w'\hat{y}_t\right)^2 / (w \hat{H}_t^{(m)} w), \tag{33}$$

where w is the vector of the EW portfolio weights. The definition of the losses for MV portfolio returns and MV* portfolio returns, denoted as $MVSCORE$ and $MVSCORE^*$, respectively, is analog, with the caveat that in these cases, the vector of the portfolio weights is the appropriate function of $\hat{H}_t^{(m)}$. Given any of such score-based losses, under Gaussianity $\bar{d}^{(m)}$ is minimum in large samples if and only if $\hat{H}_t^{(m)} = H_t^0$.

□ **Diebold-Mariano losses.** The Amisano-Giacomini score-based loss accounts for misspecifications in either the model of H_t, or the asset conditional distribution, or both. A loss function that accounts for misspecification in the model of H_t only is the Diebold-Mariano MSE-based loss. Again we consider a loss for the variance process and a loss for the correlation process. The variance loss is

$$d_t^{(m)} = VARMSE_t^{(m)} = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{y}_{i,t}^2 - \hat{h}_{i,t}^{(m)} \right)^2 , \tag{34}$$

and the correlation loss is

$$d_t^{(m)} = CORRMSE_t^{(m)} = \frac{1}{N(N-1)/2} \sum_{i<j=2,\ldots,N} \left(\hat{\varepsilon}_{i,t}^{(m)} \hat{\varepsilon}_{j,t}^{(m)} - \hat{\rho}_{ij,t}^{(m)} \right)^2 , \tag{35}$$

where $\hat{\rho}_{ij,t}^{(m)}$ is the out-of-sample estimate of $\rho_{ij,t}$. The correlation loss depends on the estimated variances through $\hat{\varepsilon}_t^{(m)}$. The portfolio MSE-based loss is defined as

$$d_t^{(m)} = EWMSE_t^{(m)} = \left(\left(\hat{w}_t' \hat{y}_t \right)^2 - \hat{w}_t' \hat{H}_t^{(m)} \hat{w}_t \right)^2 . \tag{36}$$

The definition of the losses for MV portfolio returns and MV* portfolio returns, denoted as $MVMSE$ and $MVMSE^*$, respectively, is analog. Given any of such MSE-based losses, $\bar{d}^{(m)}$ is minimum in large samples if and only if $\hat{H}_t^{(m)} = H_t^0$.

The previously described loss functions requires as an input the true and unknown variances or correlations. Those have been replaced by quantities measured using the observed returns. Such a choice is clearly suboptimal since the proxy we are using is contaminated by a noise. However, as discussed in Patton and Sheppard (2009), the previous loss functions are robust to the noise of the volatility proxy used. Alternatively, other approaches could have been considered, such as the use of realized variances as in Laurent et al. (2010), but that require the availability of high frequency data.

5.2.2 Regression-based Specification Tests

□ **Engle-Colacito regression.** Consider the regression model $\{(\hat{w}_t' \hat{y}_t)^2 / (\hat{w}_t' \hat{H}_t^{(m)} \hat{w}_t)\} - 1 = \gamma + \xi_t$, where ξ_t is an innovation term. By construction, γ is zero if $\hat{H}_t^{(m)} = H_t^0$. Hence, the significance of a test of $\gamma = 0$, where HAC-robust standard errors are required, provides evidence of misspecification.

□ **Dynamic quantile test.** The $\alpha \times 100\%$ Value-at-Risk (VaR) at time t for a given portfolio, $w_t^{0\prime} y_t$, is denoted as $VaR_t^0(\alpha)$ and is defined as the $\alpha \times 100\%$-quantile of the conditional distribution of $w_t^{0\prime} y_t$. Under Gaussianity it holds $VaR_t^0(\alpha) = \Phi^{-1}(\alpha)\sqrt{w_t^{0\prime} H_t^0 w_t^0}$, where $\Phi(\cdot)$ is the standard Gaussian distribution function. Define $HIT_t = 1$ if $\hat{w}_t' \hat{y}_t < \widehat{VaR}_t(\alpha)$ and $HIT_t = 0$ otherwise, where $\widehat{VaR}_t(\alpha)$ is the estimated VaR. The dynamic quantile test is an F-test of the hypothesis that all coefficients as well as the intercept are zero in a regression of $\{HIT_t - \alpha\}$ on its past, the current VaR and any other variables. In this paper, fifteen lags and the current VaR are used. Rejecting the null provides evidence of model misspecification.

□ **Portfolio ARCH effect test.** If the model is correctly specified, the series of the square standardized portfolio returns, $(w_t^{0\prime} y_t)^2/(w_t^{0\prime} H_t^0 w_t^0)$, is serially uncorrelated, or $w_t^{0\prime} y_t$ exhibits no ARCH effects. Rejecting the null of no serial correlation of $(\hat{w}_t' \hat{y}_t)^2/(\hat{w}_t' \hat{H}_t^{(m)} \hat{w}_t)$, then show evidence of model misspecification. We fit the test using fifteen lags on the squared residuals.[9]

5.3 Results

5.3.1 Large Dataset

Figure 6 reports the univariate estimates of the GARCH(1,1) dynamic parameters computed on the whole sample period. The figure also shows the contour plots of the estimated best GM model. The estimated number of clusters is $\hat{K} = 3$. Accordingly, most of the rolling window estimates of K oscillate between three and five with the large dataset (see Figure 7). In order to limit the calculations, the BIC-selection procedure is run only for $K = 1, 2, 3, 5, 8, 13$ rather than for $K = 1, \ldots, N$, which does not cause any notable loss of generality, as shown by the fact that the rolling window estimate of K is almost always less than 13 (see Figure 7).

Table 2 refers to the Amisano-Giacomini tests computed for the joint VC-cDCC estimator. Our particular focus is on the comparison of the VC-cDCC estimators with data driven partition, denoted as $cDCC_{VC}$ and $cDCC_{VC}^{Trg}$, with the traditional cDCC estimator. We first discuss the results relative to the targeted VC-cDCC estimators (see the left column of Table 2). The $VARSCORE$ table shows that allowing for variance clustering does not provide sensible improvements with respect to the cDCC estimator. The test statistic comparing $cDCC_{VC}^{Trg}$ and $cDCC$

[9]The actual size of all considered regression-based tests can be different from the nominal size because of the estimation error due to replacing H_t^0 with $\hat{H}_t^{(m)}$.

is equal to $-.61$, which is in favor of $cDCC_{VC}^{Trg}$ (the sign is negative) but is not significant at the 5% level. The $CORRSCORE$ table shows different results. The correlation predictions are sensibly improved by allowing for variance clustering (the last column of the table is all negative and strongly significant). Among the considered VC-cDCC estimators, $cDCC_{VC}^{Trg}$ provides the second best performance against $cDCC$, with a test statistic equal to -6.69 (the best performance is provided by $cDCC_3$ with a test statistic equal to -6.72). The good performance of the *a priori* partitioned estimator, $cDCC_*^{Trg}$, is a fictitious effect. Note, in fact, that the overall best estimator is the one-cluster estimator, $cDCC_1^{Trg}$, as shown in the $CORRSCORE$ table (the elements in the first row are all negative and significant), suggesting that the *a priori* sector partition basically acts as a random partition. Since there are only two clusters in the *a priori* partition, the loss of efficiency of $cDCC_*^{Trg}$ with respect to the well-specified estimator $cDCC_1^{Trg}$ is small. This conjecture is supported by the fact that the sign of most pairwise comparison tests between $cDCC_*^{Trg}$ and $cDCC_1^{Trg}$ is in favor of $cDCC_1^{Trg}$. As for the results from the portfolio-based pairwise comparisons (see $EVSCORE$, $MVSCORE$ and $MVSCORE^*$ in Table 2), the test to $cDCC_{VC}^{Trg}$ and $cDCC$, though not significant, is in favor of $cDCC_{VC}^{Trg}$. The results relative to the non-targeted VC-cDCC estimators (see the right column of Table 2) indicate that the overall message is that the performances are worse than those of the corresponding targeted versions. These results are in accordance with the simulation evidence discussed in the previous sections. Table A.1 in the Appendix reports the Amisano-Giacomini pairwise comparison tests computed for the sequential VC-cDCC estimators. In accordance with the simulation evidence, the pattern of the results is very similar to that obtained with the joint estimators.

The results from the Diebold-Mariano pairwise comparison tests (see Table 3 for an example) are analogous to — though less decisive than — those from the Amisano-Giacomini tests. In particular, it is confirmed that the performance of the sequential estimators are very similar to that of the corresponding joint estimators, and that the targeted estimators perform better than the corresponding non-targeted estimators. Provided that the targeting is applied, the prediction of the correlation process is significantly improved by allowing for variance clustering (see the last column of the left $CORRSCORE$ table). As desired, $cDCC_{VC}^{Trg}$ significantly outperforms $cDCC$. Note that the *a priori* partitioned estimator, $cDCC_*^{Trg}$, is significantly better than $cDCC$, but it is also significantly worse than the one-cluster estimator, $cDCC_1^{Trg}$. This difference supports the conjecture that the good performances of the sector partition are fictitious. As for the results from the regression-based tests (Table 4 includes some examples), note that in all tests the performances of the targeted VC-cDCC estimators are either equal or better than those of the cDCC estimator.

In both targeted and non-targeted cases, the overall message from the pairwise comparison tests is that the estimators with fixed $K \leq 3$ provide the best performances (most of the dark grey cells are concentrated on the rows related to such estimators). Since the rolling window estimates of the VC-cDCC estimator are mostly larger than three (see Figure 6), they seem to suggest that the performances of the VC-cDCC estimator could be sensibly improved by correcting it for the overestimation of the number of clusters already shown via simulation (see Figure 3).

5.3.2 Small Dataset

The univariate estimates of the GARCH(1,1) dynamic parameters for the whole sample period are plotted in Figure 6. The estimated number of clusters is $\hat{K} = 2$, and, accordingly, most of the rolling window estimates of K are equal to two (see Figure 7). In order to speed up the calculations, the BIC-selection procedure is run for $K = 1, 2, 3, 5, 8$. With respect to the large dataset, the superiority of the targeted estimators with respect to the corresponding non-targeted versions appears to be less marked, which is in accordance with the simulation results reported in section 4.2. According to all the considered comparison criteria, including the regression-based tests, all the considered VC-cDCC estimators, either joint or sequential, as well as targeted or non-targeted, perform equal to or better than the cDCC estimator. In particular, in the ten considered pairwise comparison tests, $cDCC_{VC}^{Trg}$ significantly beats $cDCC$ three times in the sequential version and four times in the joint version; $cDCC_{VC}$ significantly beats $cDCC$ five times in the sequential version and six times in the joint version.

As in the case of the large datasets, the estimators with a fixed, small number of clusters perform generally better than the remaining estimators. Furthermore, most statistically significant comparisons are related to the estimators with fixed $K \leq 2$, which seems to confirm that the VC-cDCC estimator can be improved by correcting it for the overestimation of the number of clusters.

6 Conclusions

We propose an estimation methodology to improve the performances of the traditional cDCC estimator for large systems by allowing for assets sharing the same dynamic variance parameters. The suggested estimator is fully frequentist, easy to implement, and endowed with good statistical properties. We first cluster the assets into groups, and then we estimate the model parameters subject to constant within-group variance dynamic parameters. The adopted clustering algorithm is new. It exploits the asymptotic properties of the univariate GARCH quasi-maximum-likelihood estimator, and thus it is distribution-free. Furthermore, in contrast to other clustering methodologies, the algorithm leads to the choice of the optimal number of clusters. Thanks to the equality parameter constraints induced by the estimated variance clustering structure, we are capable of jointly estimating the whole set of model dynamic parameters. Simulations and applications to real data are included, which show that the suggested estimator allows some improvements in terms of efficiency with respect to the traditional cDCC estimator, where the variance dynamics are estimated one at a time and separately from the correlation dynamics.

References

[1] Aielli, G. P. (2006), "Consistent estimation of large scale dynamic conditional correlations." Unpublished paper.

[2] Aielli, G. P. (2011), "Dynamic conditional correlation MGARCH models: on properties and estimation." Available at SSRN.

[3] Amisano, G., and Giacomini, R. (2007), "Comparing density forecasts via weighted likelihood ratio tests", *Journal of Business and Economic Statistics*, **25**, pp. 177-190.

[4] Bauwens L. and Rombouts, J. V. K. (2007), "Bayesian Clustering of Many Garch Models", *Econometric Reviews*, **26**, pp. 365-386.

[5] Billio, M., Caporin, M., and Gobbo, M. (2006), "Flexible Dynamic Conditional Correlation Multivariate GARCH for Asset Allocation", *Applied Financial Economics Letters*, **2**, pp. 123-130.

[6] Billio M. and M. Caporin (2009), "A generalised Dynamic Conditional Correlation model for portfolio risk evaluation", *Mathematics and Computers in Simulation*, **79**/8, pp. 2566-2578.

[7] Bollerlsev, T., (1986), "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, **31**, 307-327.

[8] Bollerlsev, T. (1990), "Modelling the Coherence in Short Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model", *The Review of Economics and Statistics*, **72**, pp. 498505.

[9] Bollerslev, T. and Wooldridge, J.M., (1992) "Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances", *Econometric Reviews*, **11**, pp. 143-172. of Paris 7.

[10] Brownlees, C., T., "On the Relation Between Firm Characteristics and Volatility Dynamics with an Application to the 2007-2009 Financial Crisis". Available at SSRN: http://ssrn.com/abstract=1695649

[11] Cappiello, L., Engle R.F., and Sheppard, K. (2006), "Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns", *Journal of Financial Econometrics*, **4**, pp. 537-572.

[12] Diebold, F.X., and Mariano, R.S., (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, **13**, pp. 253-263.

[13] Elie, L. and Jeantheau, T. (1995), "Estimation in conditionally heteroskedastic models", C.R. Acad. Sci. Paris Sér. I **320** (1995) pp. 1255-1258.

[14] Engle, R.F. (2002), "Dynamic Conditional Correlation — a Simple Class of Multivariate GARCH Models", *Journal of Business and Economic Statistics*, **20**, pp. 339-350.

[15] Engle, R.F., and Colacito, R., (2006), "Testing and valuing dynamic correlations for asset al location", *Journal of Business and Economic Statistics*, **24**, pp. 238-253.

[16] Engle, R. F., and Kelly, B. T., (2009), "Dynamic Equicorrelation". NYU Working Paper No. FIN-08-038

[17] Engle, R.F. and Manganelli, S. (2001), "CAViaR: Conditional Value at Risk by Regression Quantiles", NBER Working Paper 7341.

[18] Engle, R.F., Shephard, N. and Sheppard, K. (2009), "Fitting Vast Dimensional Time-Varying Covariance Models", Working Paper FIN-08-009, NYU Stern School of Business, Department of Finance.

[19] Engle, R.F., and Sheppard, K. (2001), "Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH", Working Paper FIN-01-027, NYU Stern School of Business, Department of Finance.

[20] Francq, C., Horvath, L. and Zakoïan, J-M. (2009), "Merits and Drawbacks of Variance Targeting in GARCH Models," Working Papers 2009-17, Centre de Recherche en Economie et Statistique.

[21] Franses, P.H. and C.M. Hafner, C.M. (2009), "A Generalised Dynamic Conditional Correlation Model: Simulation and Application to Many Assets", *Econometric Reviews*, 28-6, 612-631.

[22] Hafner, C.M., and O. Reznikova (2010), "On the estimation of dynamic conditional correlation models", *Computational Statistics and Data Analysis*, forthcoming.

[23] Laurent, S., Rombouts, J.V.K., and Violante, F., 2010, "On the forecasting accuracy of multivariate GARCH models",*Journal of Applied Econometrics*, forthcoming.

[24] Liao, T.W. (2005), "Clustering of time series data — a survey", *Pattern Recognition*, 38, 1857-1874.

[25] McAleer, M., Chan, F., Hoti, S., and Lieberman, O. (2008), "Generalized Autoregressive Conditional Correlation", *Econometric Theory*, **24**, 1554-1583.

[26] Otranto, E., (2008), "Clustering heteroskedastic time series by model-based procedures", *Computational Statistics and Data Analysis*, **52-10**, 4685-4698.

[27] Pakel, C., Shephard, N., and Sheppard, K. (2011), "Nuisance parameters, composite likelihoods, and a panel of GARCH models", *Statistica Sinica*, **21**, 307-329.

[28] Paolella, M. S. (2010), "ALRIGHT: Asymmetric LaRge-Scale (I)GARCH with Hetero-Tails", Swiss Finance Institute Research Paper Series N10 27, available at SSRN, http://ssrn.com/abstract=1628146.

[29] Patton, A.J., and Sheppard, K., (2009), "Evaluating volatility and correlation forecasts", in Andersen, T.G., Davis, R.A., Kreiß, J.P., and Mikosch, T., (eds.), *Handbook of Financial Time Series*, Springer.

[30] Pesaran, M. H. and Pesaran, B. (2007), "Modelling volatilities and conditional correlations in futures markets with a multivariate t distribution", Unpublished paper: Department of Economics, University of Cambridge.

[31] Severini, T. A. (1998), "Likelihood functions for inference in the presence of a nuisance parameter", *Biometrika*, **85**, pp. 507-522.

[32] Wooldridge, J.M. (1994), "Estimation and Inference for Dependent Processes", in *Handbook of Econometrics* (Vol. IV) , eds. R. Engle and D. McFadden, New York: Elsevier Science, pp. 2639–2738.

[33] Xu, R. and Wunsch, II, D.C. (2009), Clustering, IEEE Press Series on Computational Intelligence, IEEE Press and Wiley, NJ.

Underlying variance clustering structure

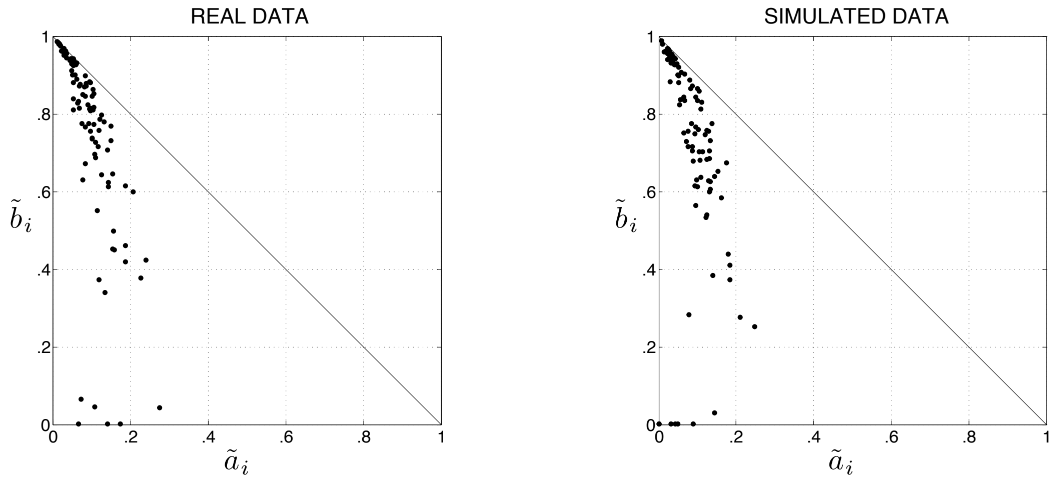

Figure 1: On the left: plot of estimates of univariate GARCH(1,1) dynamic parameters from $N = 100$ series of real data. The dataset is the equity dataset described in section 5.1. On the right: plot of estimates of univariate GARCH(1,1) dynamic parameters from simulated data. The DGP underlying the simulated series provides for $K^0 = 4$ distinct GARCH dynamics, namely, $\mu_1^0 = (.13, .56)$ for $N_1 = 13$ assets, $\mu_2^0 = (.11, .73)$ for $N_2 = 39$ assets, $\mu_3^0 = (.03, .96)$ for $N_3 = 42$ assets, and $\mu_4^0 = (.14, 0)$ for $N_4 = 6$ assets. The simulated correlation process is set as constant and equal to $s_{ij} = s = .3$ for all asset pairs.

GARCH-GM clustering

$N = 100$

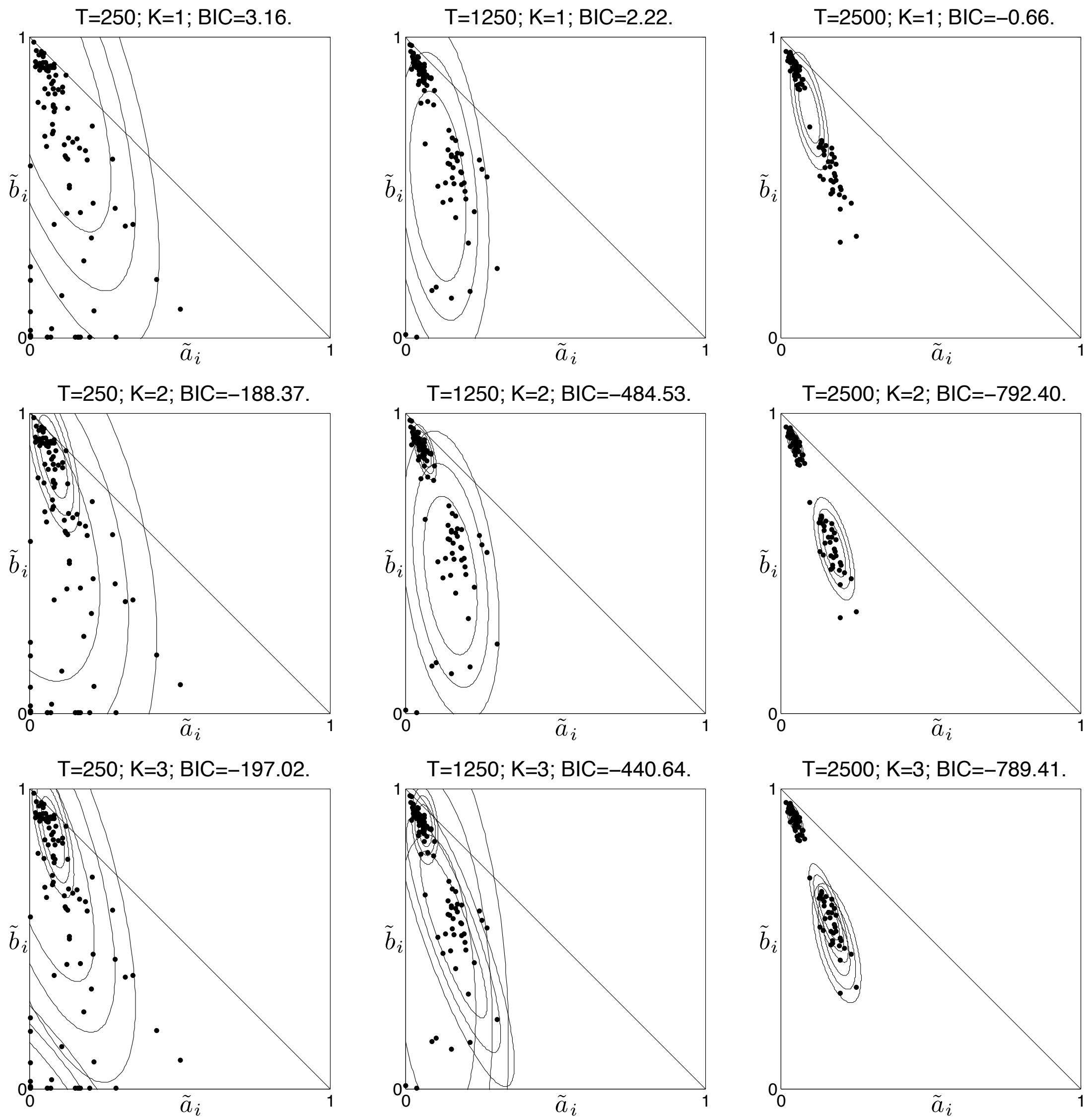

Figure 2: Univariate QML GARCH(1,1) dynamic parameter estimates from simulated data and related GARCH-GM component contour plots. The DGP variance parameters are set as in Table 1 - Large dataset, with constant conditional correlations set as $s_{ij}^0 = s = .3$ for all asset pairs. The true number of clusters is $K^0 = 2$. The component contour plots are drawn in correspondence to the Gaussian confidence sets of level $.50, .75, .90$. The lower triangle of each plot is the admissible parameter space, $\{(a, b) : a \geq 0, b \geq 0, a + b < 1\}$.

Clustering estimation performances

$N = 100$

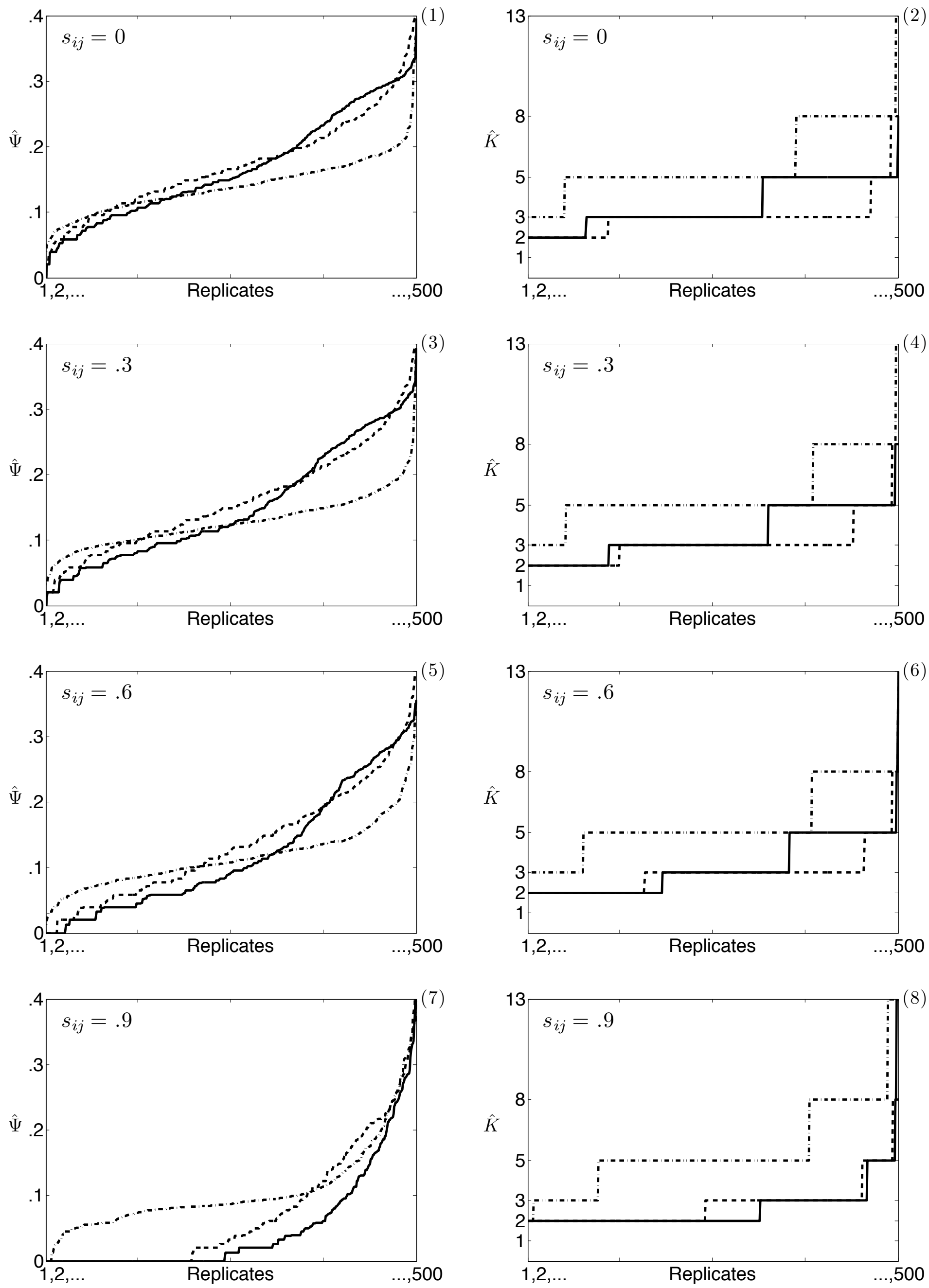

Figure 3: Percentage of incorrect allocations, $\hat{\Psi}$, and of estimated number of clusters, $\hat{K}$, arranged in ascending order. Straight lines for the GARCH-GM clustering; dashed lines for the HE-GM clustering, and dash-dotted lines for the HO-GM clustering. The true number of clusters is $K^0 = 2$.

Clustering estimation performances

$N = 10$

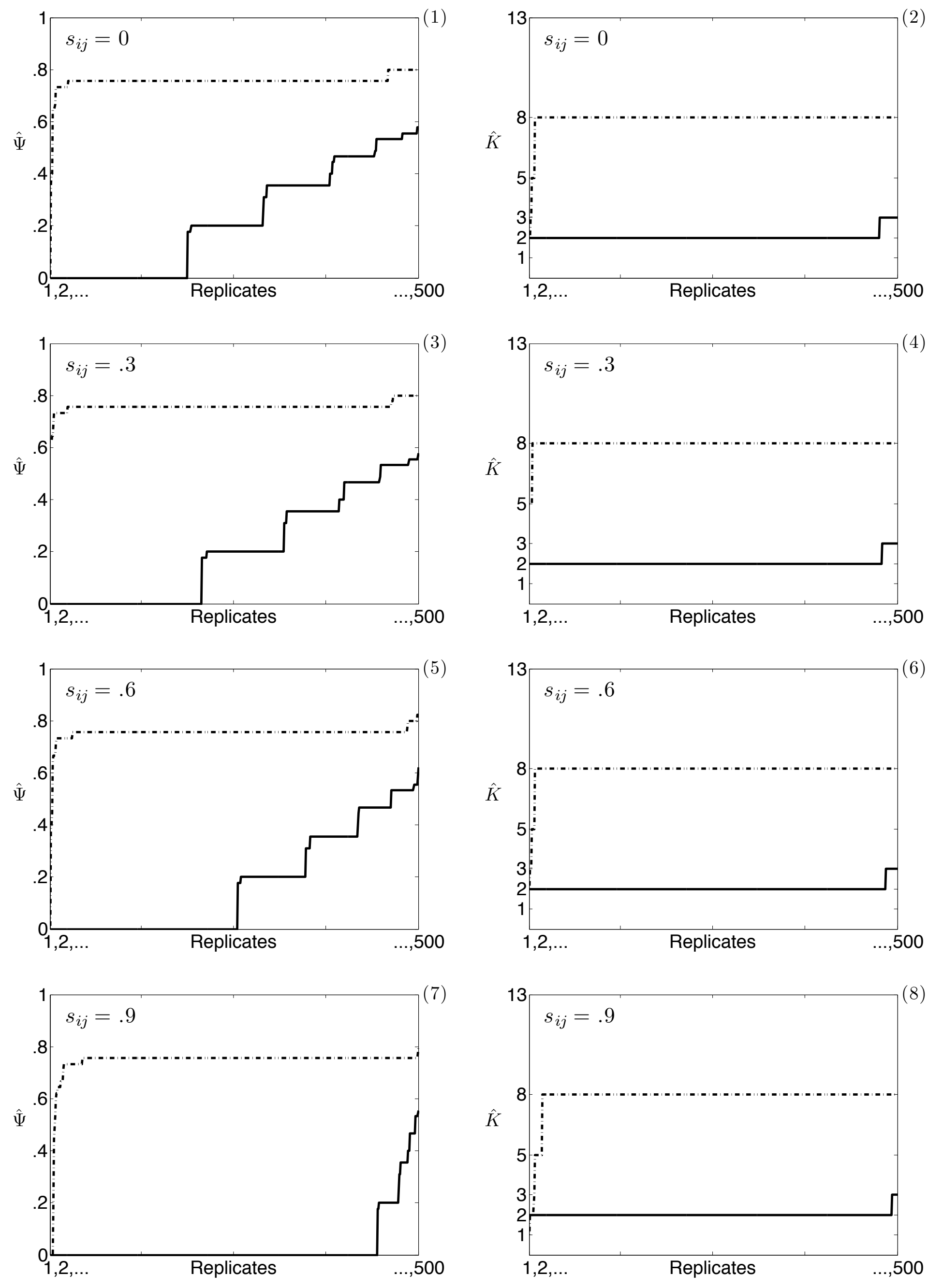

Figure 4: Percentage of incorrect allocations, $\hat{\Psi}$, and estimated number of clusters, $\hat{K}$, arranged in ascending order. Straight lines for the GARCH-GM clustering; dash-dotted lines for the HO-GM clustering. The true number of clusters is $K^0 = 2$. The due to large number of parameters, the HE-GM clustering is infeasible for $N = 10$.

Variance/correlation estimation performances

$N = 100$

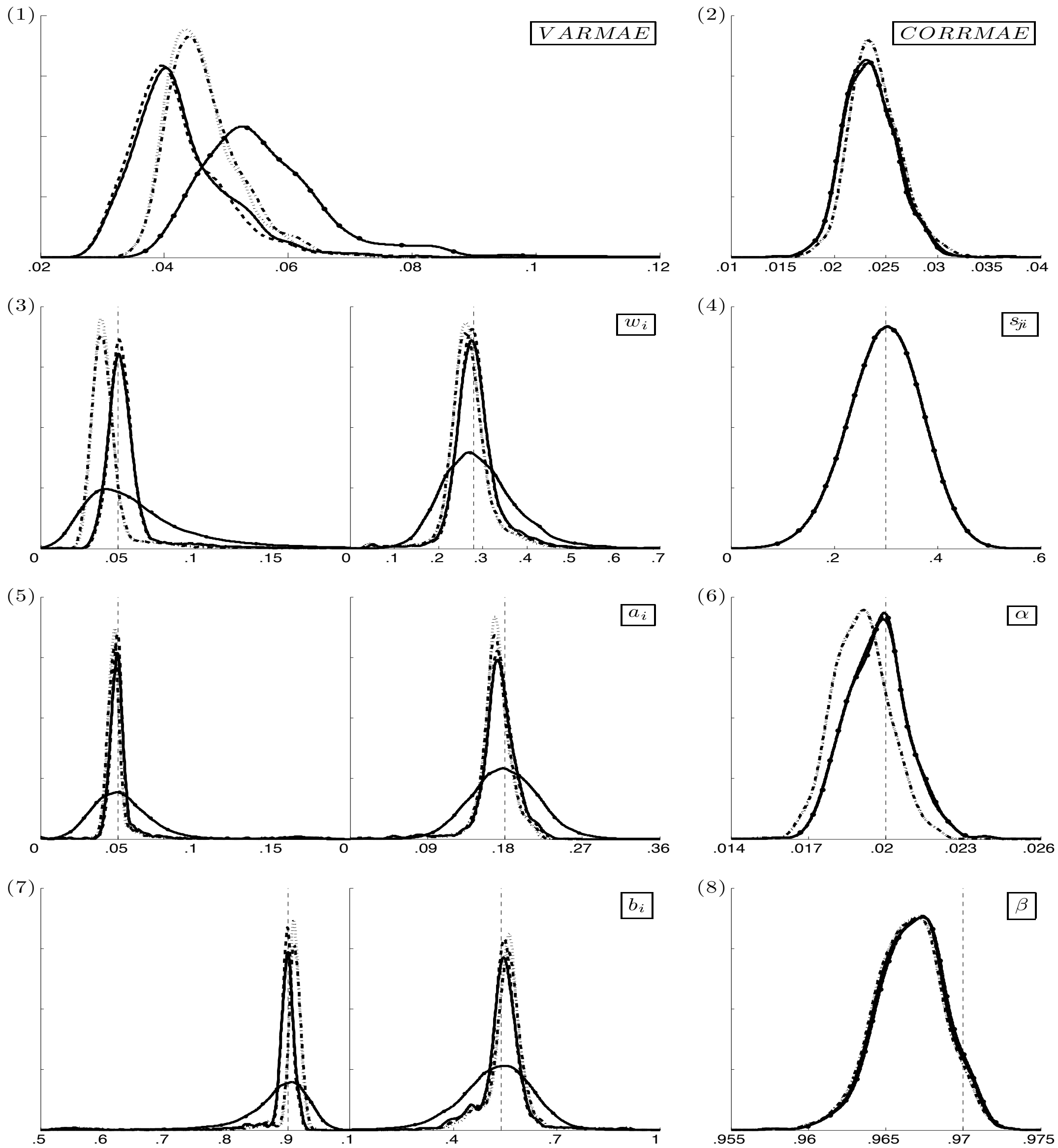

Figure 5: cDCC estimator in straight dotted lines; targeted joint VC-cDCC estimator in straight lines; targeted sequential VC-cDCC estimator in dashed lines; non-targeted joint VC-cDCC estimator in dashed dotted lines; non-targeted sequential VC-cDCC estimator in dotted lines.

Estimates of GARCH(1,1) dynamic parameters from real data

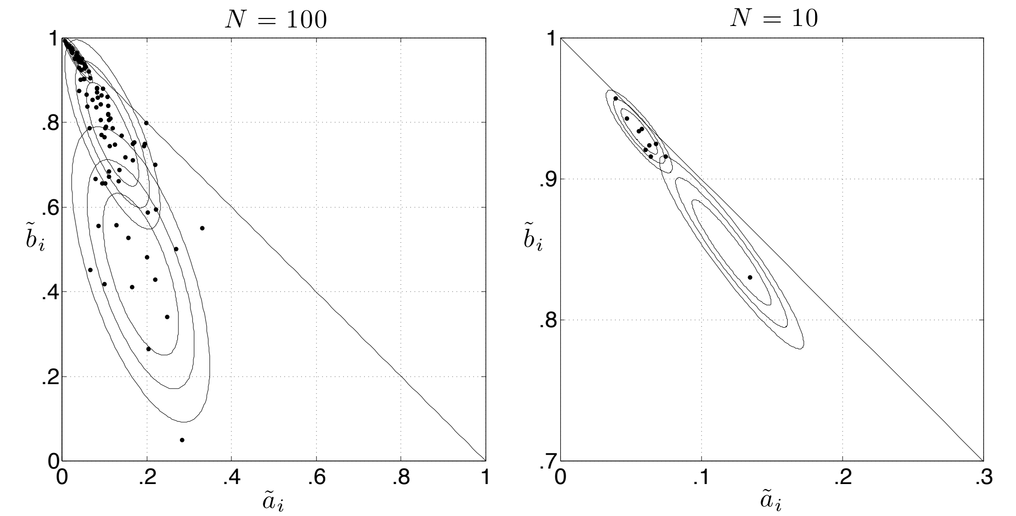

Figure 6: Univariate QML estimates of GARCH(1,1) dynamic parameters obtained from real data. The contour plots are relative to the components of the estimated GM-GARCH model ($\hat{K} = 3$ on the left; $\hat{K} = 2$ on the right). The component contour plots are drawn in correspondence to the Gaussian confidence sets of level $.50, .75, .90$.

Estimated number of clusters from real data

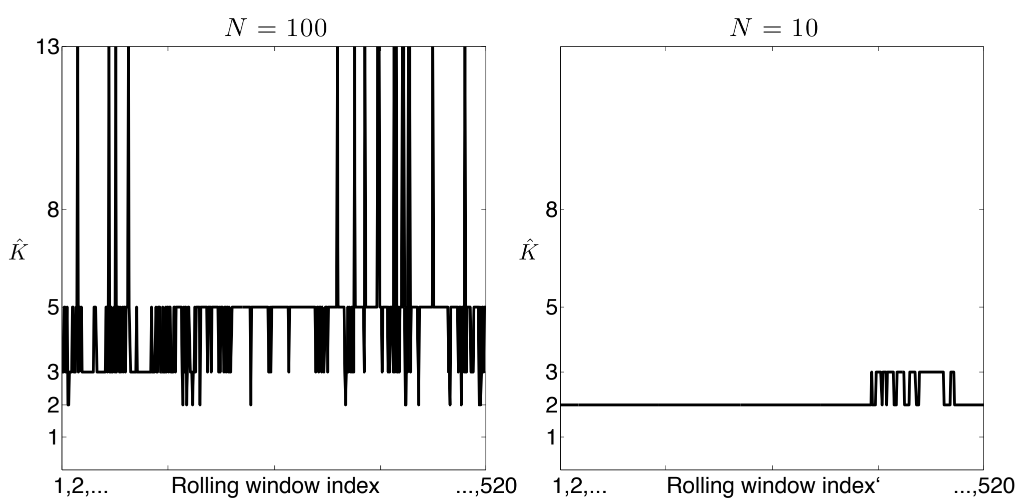

Figure 7: Rolling window estimates of the number of clusters.

Large dataset ($N = 100$)		Small dataset ($N = 10$)	
$K^0 = 2$		$K^0 = 2$	
$N_1^0 = 66$, $N_2^0 = 34$		$N_1^0 = 9$, $N_2^0 = 1$	
$\tau_i^0 = 1,$	$i = 1, \dots, N$	$\tau_i^0 = 1$	$i = 1, \dots, N$
$(a_i^0, b_i^0) = (.05, .90),$	$i = 1, \dots, N_1^0$	$(a_i^0, b_i^0) = (.06, .93),$	$i = 1, \dots, N_1^0$
$(a_i^0, b_i^0) = (.18, .54),$	$i = N_1^0 + 1, \dots, N_1^0 + N_2^0$	$(a_i^0, b_i^0) = (.12, .85),$	$i = N_1^0 + 1 = N_1^0 + N_2^0$

Table 1: DGP variance parameter values.

Amisano-Giacomini pairwise comparison tests

$N = 100$; joint dynamic parameter estimation.

$VARSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.77	0.54	0.49	-1.42	-0.38	-0.53
$cDCC_2^{Trg}$	–	–	0.13	-0.04	-1.61	-1.29	-1.06
$cDCC_3^{Trg}$	–	–	–	-0.30	-1.03	-2.54	-1.50
$cDCC_5^{Trg}$	–	–	–	–	-1.04	-2.75	-1.39
$cDCC_*^{Trg}$	–	–	–	–	–	-0.02	-0.35
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-0.61
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.76	0.29	0.03	-1.28	-1.05	-0.02
$cDCC_2$	–	–	-0.34	-0.45	-1.70	-1.98	-0.48
$cDCC_3$	–	–	–	-0.37	-1.10	-2.71	-0.37
$cDCC_5$	–	–	–	–	-0.56	-2.17	-0.11
$cDCC_*$	–	–	–	–	–	-0.41	0.51
$cDCC_{VC}$	–	–	–	–	–	–	1.75
$cDCC$	–	–	–	–	–	–	–

$CORRSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-2.45	-2.89	-3.96	-3.15	-3.96	-5.84
$cDCC_2^{Trg}$	–	–	-2.32	-4.20	0.87	-4.19	-6.53
$cDCC_3^{Trg}$	–	–	–	-1.39	2.24	-2.74	-6.72
$cDCC_5^{Trg}$	–	–	–	–	3.37	-1.81	-6.66
$cDCC_*^{Trg}$	–	–	–	–	–	-3.48	-6.00
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-6.69
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	2.36	1.48	1.49	-1.05	0.87	0.27
$cDCC_2$	–	–	-0.69	0.24	-3.37	-0.91	-1.21
$cDCC_3$	–	–	–	0.92	-2.38	-0.71	-1.10
$cDCC_5$	–	–	–	–	-2.01	-2.57	-3.07
$cDCC_*$	–	–	–	–	–	1.49	0.70
$cDCC_{VC}$	–	–	–	–	–	–	-1.08
$cDCC$	–	–	–	–	–	–	–

$EWSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	1.64	1.70	1.04	-0.88	1.05	0.26
$cDCC_2^{Trg}$	–	–	1.36	0.06	-2.73	0.06	-0.70
$cDCC_3^{Trg}$	–	–	–	-2.04	-2.06	-2.06	-2.12
$cDCC_5^{Trg}$	–	–	–	–	-1.48	-0.01	-1.23
$cDCC_*^{Trg}$	–	–	–	–	–	1.50	0.52
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.21
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.82	0.18	-1.14	-0.61	-1.25	-0.69
$cDCC_2$	–	–	-1.01	-2.01	-1.90	-2.10	-1.59
$cDCC_3$	–	–	–	-2.17	-0.71	-2.33	-1.58
$cDCC_5$	–	–	–	–	0.76	-0.67	1.65
$cDCC_*$	–	–	–	–	–	-0.85	-0.33
$cDCC_{VC}$	–	–	–	–	–	–	1.81
$cDCC$	–	–	–	–	–	–	–

$MVSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.49	-0.70	0.39	0.10	-0.13	-0.20
$cDCC_2^{Trg}$	–	–	-1.18	0.05	-0.43	-0.58	-0.45
$cDCC_3^{Trg}$	–	–	–	1.28	0.71	0.75	0.18
$cDCC_5^{Trg}$	–	–	–	–	-0.35	-0.87	-0.61
$cDCC_*^{Trg}$	–	–	–	–	–	-0.14	-0.21
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-0.20
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-1.84	-2.65	-1.86	-0.91	-2.15	-1.13
$cDCC_2$	–	–	-1.66	-0.73	1.07	-1.13	-0.18
$cDCC_3$	–	–	–	1.19	2.42	0.73	1.22
$cDCC_5$	–	–	–	–	1.47	-0.69	0.51
$cDCC_*$	–	–	–	–	–	-1.84	-0.80
$cDCC_{VC}$	–	–	–	–	–	–	0.82
$cDCC$	–	–	–	–	–	–	–

$MVSCORE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.64	-0.38	-0.01	-1.44	-0.73	-0.68
$cDCC_2^{Trg}$	–	–	0.24	0.63	0.15	-0.15	-0.46
$cDCC_3^{Trg}$	–	–	–	0.54	-0.11	-0.54	-0.60
$cDCC_5^{Trg}$	–	–	–	–	-0.42	-1.44	-0.86
$cDCC_*^{Trg}$	–	–	–	–	–	-0.27	-0.46
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-0.40
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-2.05	-1.42	-0.81	-0.22	-1.57	-0.95
$cDCC_2$	–	–	0.54	1.16	1.90	0.31	0.19
$cDCC_3$	–	–	–	0.96	1.42	-0.30	-0.16
$cDCC_5$	–	–	–	–	0.75	-1.35	-0.68
$cDCC_*$	–	–	–	–	–	-1.55	-0.95
$cDCC_{VC}$	–	–	–	–	–	–	0.03
$cDCC$	–	–	–	–	–	–	–

Table 2: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Diebold-Mariano pairwise comparison tests

$N = 100$; joint dynamic parameter estimation.

$VARMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.98	-0.65	-1.64	0.81	-1.81	-1.52
$cDCC_2^{Trg}$	–	–	0.02	-1.53	1.20	-1.75	-1.15
$cDCC_3^{Trg}$	–	–	–	-2.28	0.84	-2.73	-1.20
$cDCC_5^{Trg}$	–	–	–	–	1.63	-0.91	0.20
$cDCC_*^{Trg}$	–	–	–	–	–	-1.79	-1.72
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	0.47
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-2.40	-1.35	-2.00	-1.78	-2.20	-1.49
$cDCC_2$	–	–	-0.12	-1.38	1.67	-1.63	-0.55
$cDCC_3$	–	–	–	-2.63	0.98	-3.04	-0.76
$cDCC_5$	–	–	–	–	1.70	-0.64	1.02
$cDCC_*$	–	–	–	–	–	-1.89	-1.18
$cDCC_{VC}$	–	–	–	–	–	–	1.17
$cDCC$	–	–	–	–	–	–	–

$CORRMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.84	-0.84	-1.46	-2.38	-1.71	-2.92
$cDCC_2^{Trg}$	–	–	-0.61	-1.85	-0.53	-2.12	-3.62
$cDCC_3^{Trg}$	–	–	–	-1.27	0.18	-2.98	-4.33
$cDCC_5^{Trg}$	–	–	–	–	0.80	-1.84	-3.90
$cDCC_*^{Trg}$	–	–	–	–	–	-1.19	-2.91
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-3.78
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	1.97	1.36	0.95	-1.05	0.47	0.25
$cDCC_2$	–	–	-0.16	0.05	-2.74	-1.10	-0.92
$cDCC_3$	–	–	–	0.19	-2.13	-1.83	-1.25
$cDCC_5$	–	–	–	–	-1.38	-1.68	-2.80
$cDCC_*$	–	–	–	–	–	1.09	0.69
$cDCC_{VC}$	–	–	–	–	–	–	-0.25
$cDCC$	–	–	–	–	–	–	–

$EWMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	1.47	1.63	0.93	-0.95	0.94	0.22
$cDCC_2^{Trg}$	–	–	1.42	-0.02	-2.33	0.01	-0.81
$cDCC_3^{Trg}$	–	–	–	-2.29	-1.95	-2.29	-2.42
$cDCC_5^{Trg}$	–	–	–	–	-1.32	0.15	-1.35
$cDCC_*^{Trg}$	–	–	–	–	–	1.34	0.48
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.35
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.22	-0.28	-1.52	-1.11	-1.64	-1.04
$cDCC_2$	–	–	-0.86	-1.85	-1.59	-1.91	-1.37
$cDCC_3$	–	–	–	-2.19	-0.77	-2.30	-1.46
$cDCC_5$	–	–	–	–	0.85	-0.36	2.09
$cDCC_*$	–	–	–	–	–	-0.92	-0.33
$cDCC_{VC}$	–	–	–	–	–	–	2.03
$cDCC$	–	–	–	–	–	–	–

$MVMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.07	-1.01	-0.12	-1.06	-0.40	-0.02
$cDCC_2^{Trg}$	–	–	-0.89	-0.08	-0.17	-0.32	0.02
$cDCC_3^{Trg}$	–	–	–	0.77	0.86	0.79	0.61
$cDCC_5^{Trg}$	–	–	–	–	-0.05	-0.25	0.07
$cDCC_*^{Trg}$	–	–	–	–	–	-0.21	0.10
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	0.29
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-1.48	-2.34	-1.45	-1.03	-2.15	-1.39
$cDCC_2$	–	–	-1.24	-0.37	0.48	-1.16	-0.62
$cDCC_3$	–	–	–	1.03	1.97	0.62	0.51
$cDCC_5$	–	–	–	–	0.73	-0.59	-0.36
$cDCC_*$	–	–	–	–	–	-1.84	-0.78
$cDCC_{VC}$	–	–	–	–	–	–	0.17
$cDCC$	–	–	–	–	–	–	–

$MVMSE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-1.02	0.03	0.19	-1.42	-0.67	-0.15
$cDCC_2^{Trg}$	–	–	0.78	0.85	0.56	0.45	0.84
$cDCC_3^{Trg}$	–	–	–	0.27	-0.55	-0.74	-0.17
$cDCC_5^{Trg}$	–	–	–	–	-0.62	-1.39	-0.32
$cDCC_*^{Trg}$	–	–	–	–	–	-0.05	0.38
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	0.41
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-2.32	-1.58	-1.32	-0.89	-2.15	-1.77
$cDCC_2$	–	–	0.11	0.38	1.29	-0.42	-0.77
$cDCC_3$	–	–	–	0.65	1.18	-0.62	-1.04
$cDCC_5$	–	–	–	–	0.83	-1.16	-1.45
$cDCC_*$	–	–	–	–	–	-1.80	-1.41
$cDCC_{VC}$	–	–	–	–	–	–	-0.59
$cDCC$	–	–	–	–	–	–	–

Table 3: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Regression-based specification tests

$N = 100$; joint dynamic parameter estimation.

ENGLE-COLACITO REGRESSION

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	1.90	1.97	1.95	2.03	1.97	2.01	2.27
MV	3.16	3.22	3.52	3.47	3.15	3.49	4.26
MV*	1.47	1.64	1.76	1.83	1.54	1.87	2.64

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	2.37	2.28	2.23	2.28	2.38	2.29	2.27
MV	4.17	4.49	4.63	4.47	4.13	4.45	4.26
MV*	2.69	2.86	2.79	2.66	2.65	2.74	2.64

PORTFOLIO ARCH EFFECT TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	37.03	36.33	35.73	36.31	36.89	36.31	35.77
MV	17.60	16.53	17.25	16.32	18.28	19.27	21.63
MV*	32.00	36.92	26.16	26.08	32.64	27.51	31.62

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	33.53	36.64	35.98	36.67	35.04	36.06	35.77
MV	24.69	19.85	17.49	18.35	25.87	22.32	21.63
MV*	29.18	31.86	26.70	28.07	30.54	28.48	31.62

1% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	1.76	1.80	1.80	1.84	1.77	1.80	1.84
MV	0.91	0.88	1.03	0.75	0.92	1.24	2.02
MV*	1.36	1.38	1.49	1.53	1.29	1.53	2.37

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	1.85	1.85	1.84	1.85	1.84	1.85	1.84
MV	0.47	0.65	2.19	1.72	0.95	2.12	2.02
MV*	1.89	2.31	2.09	2.32	2.53	2.34	2.37

5% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	0.72	0.71	0.71	0.71	0.72	0.71	0.71
MV	1.42	1.71	0.87	0.87	1.19	1.21	2.36
MV*	1.81	1.60	1.25	1.26	1.36	1.19	1.54

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	1.11	0.71	0.97	0.97	1.10	0.97	0.71
MV	0.96	1.09	0.92	1.33	1.76	1.99	2.36
MV*	1.33	1.21	1.22	1.70	1.12	1.41	1.54

Table 4: White cells denote insignificance at 5% level, light grey cells denote significance at 5% level and dark grey cells denote significance at 1% level.

A Additional Figures and Tables

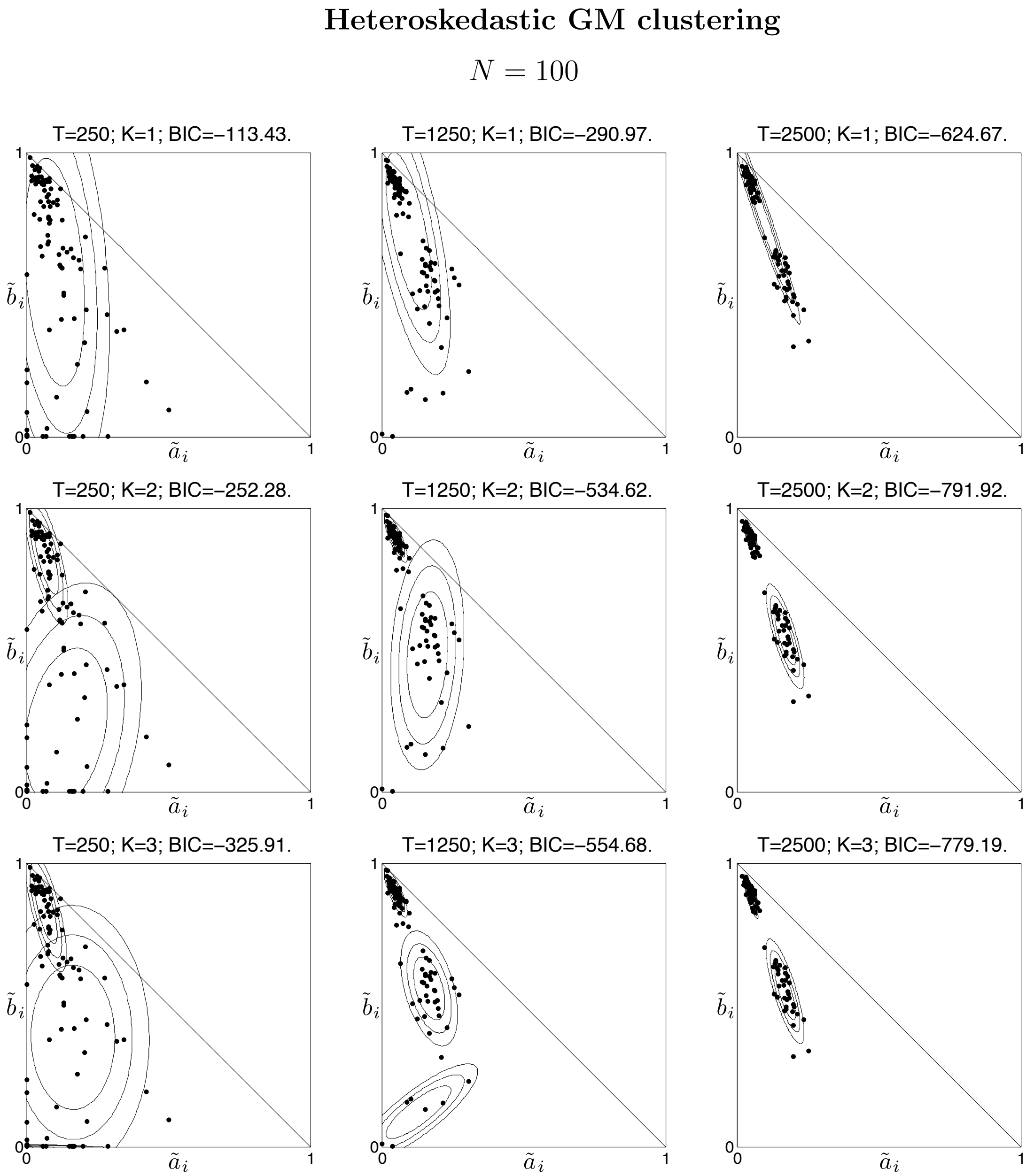

Figure A.1: Univariate QML GARCH(1,1) dynamic parameter estimates from simulated data and related HE-GM component contour plots. The DGP variance parameters are set as in Table 1 - Large dataset, with constant conditional correlations set as $s_{ij}^0 = s = .3$ for all asset pairs. The true number of clusters is $K^0 = 2$. The component contour plots are drawn in correspondence to the Gaussian confidence sets of level $.50, .75, .90$. The lower triangle of each plot is the admissible parameter space, $\{(a, b) : a \geq 0, b \geq 0, a + b < 1\}$.

Homoskedastic GM clustering

$N = 100$

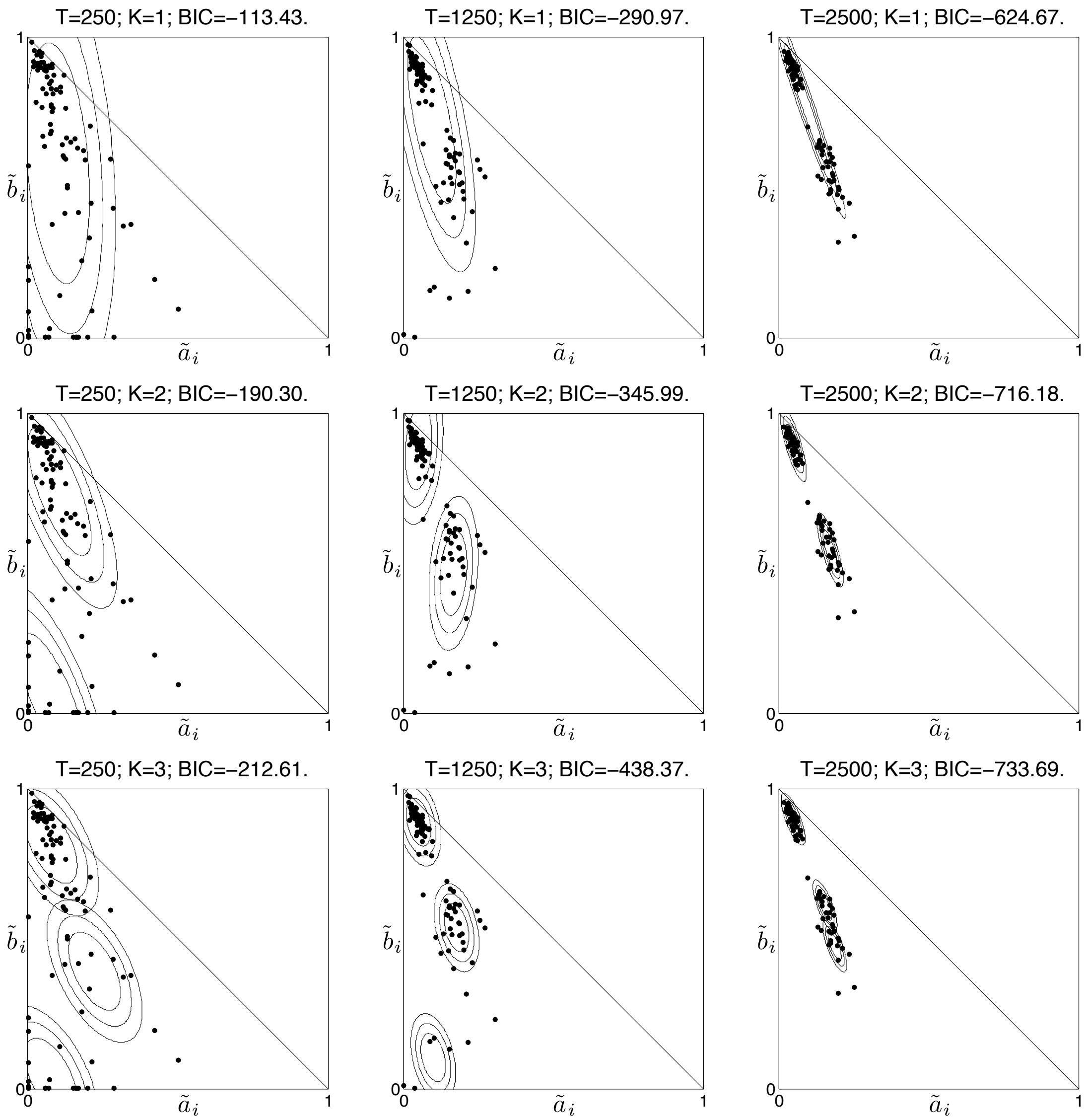

Figure A.2: Univariate QML GARCH(1,1) dynamic parameter estimates from simulated data and related HO-GM component contour plots. The DGP variance parameters are set as in Table 1 - Large dataset, with constant conditional correlations set as $s_{ij}^0 = s = .3$ for all asset pairs. The true number of clusters is $K^0 = 2$. The component contour plots are drawn in correspondence to the Gaussian confidence sets of level .50, .75, .90. The lower triangle of each plot is the admissible parameter space, $\{(a, b) : a \geq 0, b \geq 0, a + b < 1\}$.

GARCH-GM clustering

$N = 5$; no common GARCH dynamics

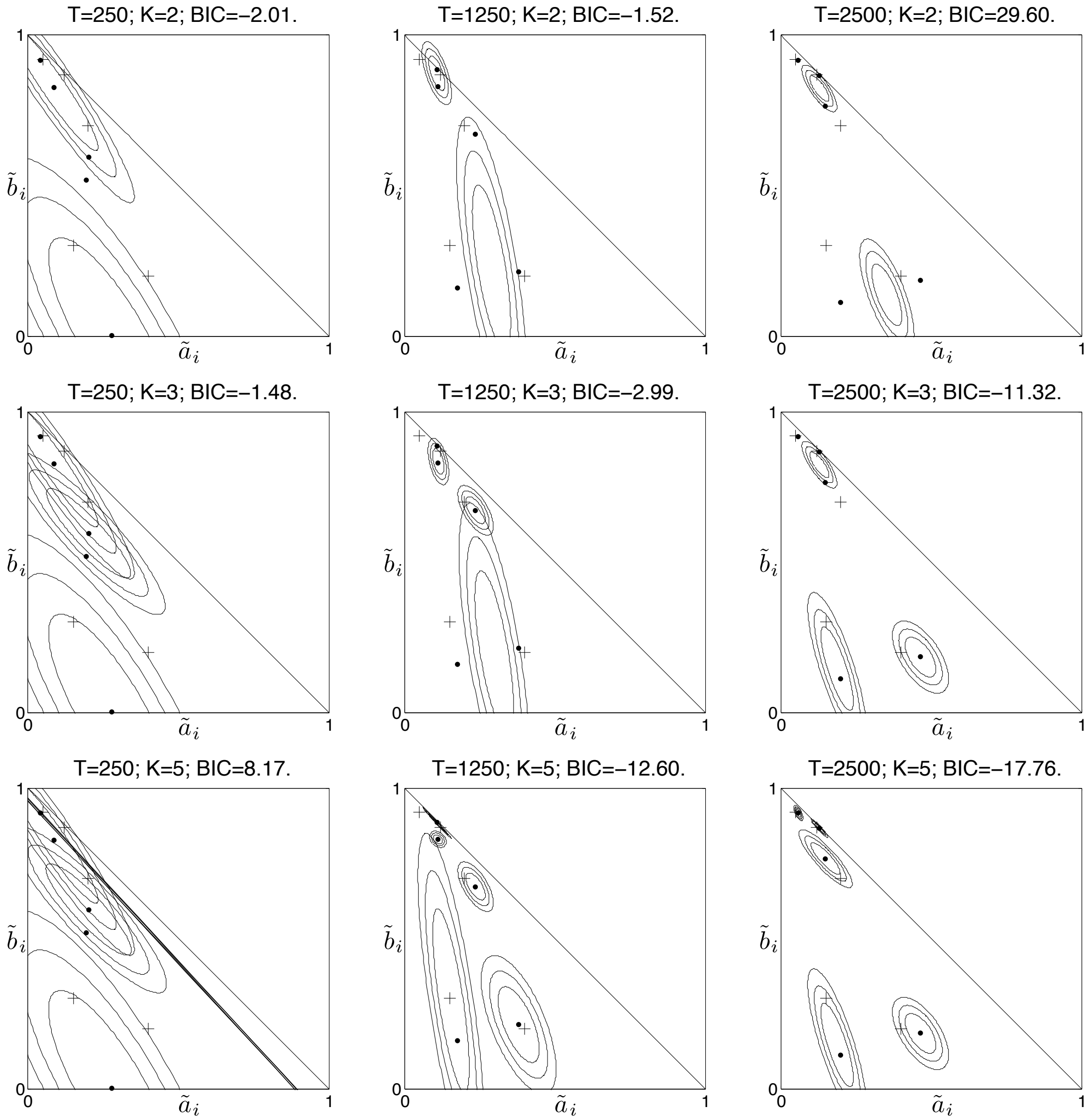

Figure A.3: Univariate QML estimates of GARCH(1,1) dynamic parameters from simulated data and related GARCH-GM component contour plots. There are $N = 5$ assets without common GARCH dynamics. The DGP variance parameter settings are $(a_1^0, b_1^0) = (.05, .92)$, $(a_2^0, b_2^0) = (.12, .87)$, $(a_3^0, b_3^0) = (.2, .7)$, $(a_4^0, b_4^0) = (.15, .3)$, $(a_5^0, b_5^0) = (.4, .2)$. The simulated correlation process is constant, with constant conditional correlations set as $s_{ij} = s = .3$ for all asset pairs. The crosses denote the true GARCH dynamics and the dots denote the univariate QML GARCH estimates. The component contour plots are drawn in correspondence to the Gaussian confidence sets of level .50, .75, .90. The lower triangle of each plot is the admissible parameter space, $\{(a, b) : a \geq 0, b \geq 0, a + b < 1\}$.

Variance/correlation estimation performances

$N = 10$

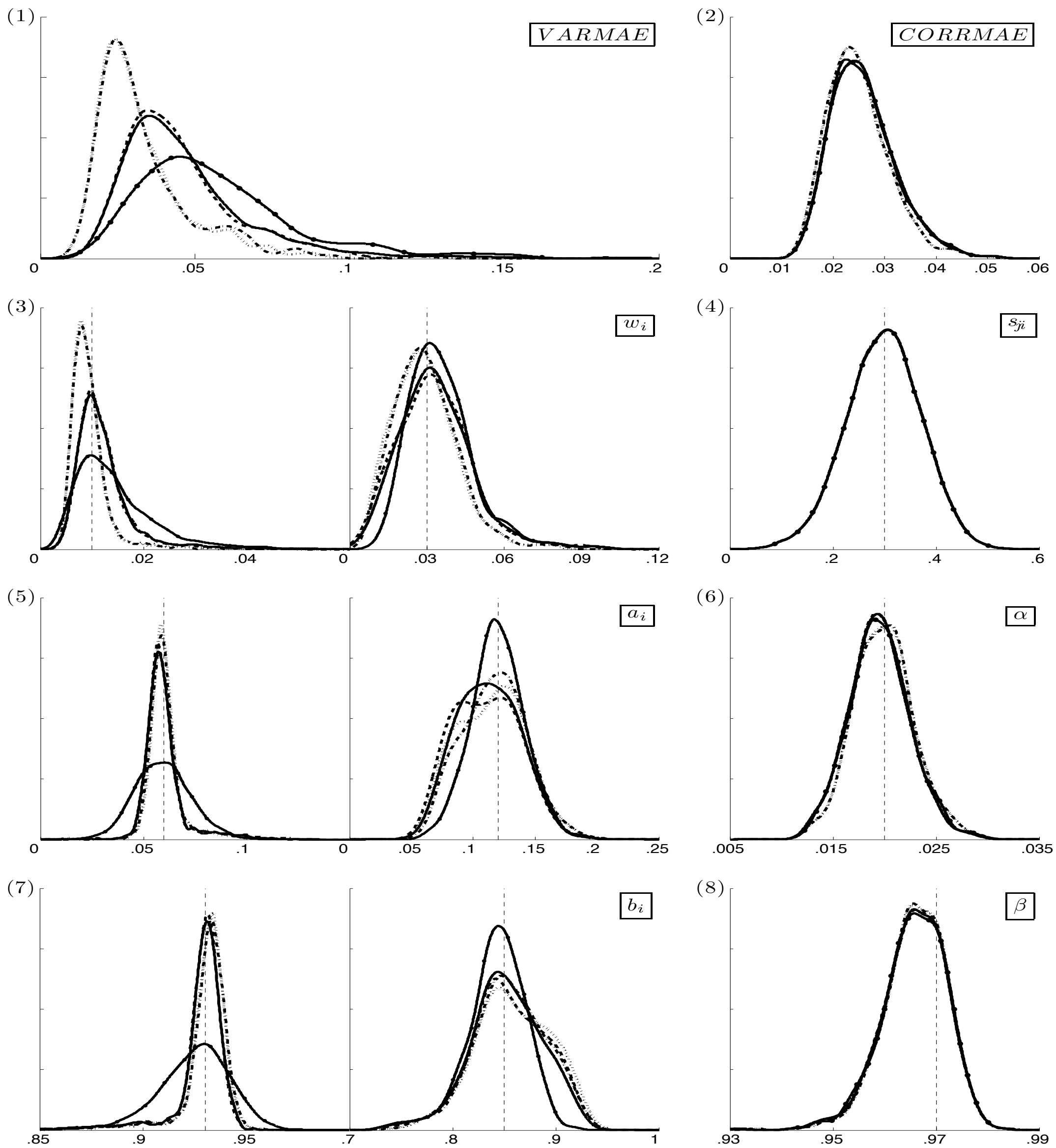

Figure A.4: cDCC estimator in straight dotted lines; targeted joint VC-cDCC estimator in straigth lines; targeted sequential VC-cDCC estimator in dashed lines; non-targeted joint VC-cDCC estimator in dashed dotted lines; non-targeted sequential VC-cDCC estimator in dotted lines.

Amisano-Giacomini pairwise comparison tests

$N = 100$; sequential dynamic parameter estimation.

$VARSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.41	0.27	0.20	-1.94	-0.64	-0.76
$cDCC_2^{Trg}$	–	–	-0.01	-0.20	-1.41	-1.56	-1.28
$cDCC_3^{Trg}$	–	–	–	-0.25	-0.91	-2.49	-1.60
$cDCC_5^{Trg}$	–	–	–	–	-0.91	-2.82	-1.54
$cDCC_*^{Trg}$	–	–	–	–	–	-0.13	-0.48
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-0.74
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.67	0.18	0.01	-0.61	-1.08	-0.11
$cDCC_2$	–	–	-0.40	-0.42	-0.98	-1.97	-0.53
$cDCC_3$	–	–	–	-0.25	-0.44	-2.57	-0.40
$cDCC_5$	–	–	–	–	-0.19	-2.15	-0.31
$cDCC_*$	–	–	–	–	–	-0.84	0.06
$cDCC_{VC}$	–	–	–	–	–	–	1.56
$cDCC$	–	–	–	–	–	–	–

$CORRSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-2.33	-3.01	-3.99	-3.41	-4.06	-5.66
$cDCC_2^{Trg}$	–	–	-2.46	-4.48	0.53	-4.54	-6.53
$cDCC_3^{Trg}$	–	–	–	-2.89	2.00	-3.46	-6.40
$cDCC_5^{Trg}$	–	–	–	–	3.19	-0.96	-6.20
$cDCC_*^{Trg}$	–	–	–	–	–	-3.34	-5.67
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-6.16
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	1.57	1.24	1.07	0.63	0.56	0.47
$cDCC_2$	–	–	0.10	0.28	-1.34	-0.50	-0.29
$cDCC_3$	–	–	–	0.36	-1.02	-1.02	-0.45
$cDCC_5$	–	–	–	–	-0.88	-2.36	-1.13
$cDCC_*$	–	–	–	–	–	0.35	0.31
$cDCC_{VC}$	–	–	–	–	–	–	0.13
$cDCC$	–	–	–	–	–	–	–

$EWSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	1.22	1.53	0.82	-1.25	0.66	-0.20
$cDCC_2^{Trg}$	–	–	1.41	-0.07	-2.43	-0.35	-1.32
$cDCC_3^{Trg}$	–	–	–	-2.12	-2.07	-2.36	-2.56
$cDCC_5^{Trg}$	–	–	–	–	-1.47	-0.63	-1.91
$cDCC_*^{Trg}$	–	–	–	–	–	1.30	0.15
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.79
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.86	0.19	-1.17	0.07	-1.40	-1.01
$cDCC_2$	–	–	-1.13	-2.01	-1.20	-2.23	-1.91
$cDCC_3$	–	–	–	-2.42	-0.18	-2.71	-2.12
$cDCC_5$	–	–	–	–	1.17	-0.89	0.41
$cDCC_*$	–	–	–	–	–	-1.38	-1.06
$cDCC_{VC}$	–	–	–	–	–	–	1.03
$cDCC$	–	–	–	–	–	–	–

$MVSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.31	-0.77	-0.40	-0.07	-0.36	-0.32
$cDCC_2^{Trg}$	–	–	-1.20	-0.78	-0.32	-0.75	-0.54
$cDCC_3^{Trg}$	–	–	–	0.24	0.76	0.41	0.07
$cDCC_5^{Trg}$	–	–	–	–	0.38	0.15	-0.10
$cDCC_*^{Trg}$	–	–	–	–	–	-0.34	-0.30
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-0.17
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-2.01	-2.53	-1.98	-0.39	-2.19	-1.02
$cDCC_2$	–	–	-1.43	-0.94	1.85	-1.06	0.24
$cDCC_3$	–	–	–	0.43	2.61	0.54	1.35
$cDCC_5$	–	–	–	–	2.02	0.01	1.18
$cDCC_*$	–	–	–	–	–	-2.24	-0.97
$cDCC_{VC}$	–	–	–	–	–	–	1.09
$cDCC$	–	–	–	–	–	–	–

$MVSCORE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.78	-0.70	-0.16	-1.20	-0.93	-0.87
$cDCC_2^{Trg}$	–	–	0.09	0.52	0.43	-0.26	-0.60
$cDCC_3^{Trg}$	–	–	–	0.66	0.35	-0.44	-0.66
$cDCC_5^{Trg}$	–	–	–	–	-0.12	-1.28	-0.95
$cDCC_*^{Trg}$	–	–	–	–	–	-0.60	-0.71
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-0.49
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-1.90	-1.41	-0.80	0.75	-1.58	-0.92
$cDCC_2$	–	–	0.55	1.07	2.25	0.24	0.17
$cDCC_3$	–	–	–	0.88	1.73	-0.34	-0.15
$cDCC_5$	–	–	–	–	1.02	-1.15	-0.59
$cDCC_*$	–	–	–	–	–	-1.86	-1.12
$cDCC_{VC}$	–	–	–	–	–	–	0.05
$cDCC$	–	–	–	–	–	–	–

Table A.1: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Diebold-Mariano pairwise comparison tests

$N = 100$; sequential dynamic parameter estimation.

$VARMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-1.48	-1.00	-1.93	0.29	-2.10	-1.98
$cDCC_2^{Trg}$	–	–	0.10	-1.47	1.36	-1.70	-1.23
$cDCC_3^{Trg}$	–	–	–	-2.30	0.99	-2.74	-1.38
$cDCC_5^{Trg}$	–	–	–	–	1.78	-0.89	0.01
$cDCC_*^{Trg}$	–	–	–	–	–	-1.94	-1.99
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	0.27
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-2.45	-1.39	-2.01	-1.64	-2.20	-1.53
$cDCC_2$	–	–	-0.14	-1.35	1.89	-1.59	-0.60
$cDCC_3$	–	–	–	-2.54	1.10	-2.96	-0.82
$cDCC_5$	–	–	–	–	1.78	-0.60	0.87
$cDCC_*$	–	–	–	–	–	-1.97	-1.30
$cDCC_{VC}$	–	–	–	–	–	–	1.02
$cDCC$	–	–	–	–	–	–	–

$CORRMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.81	-0.95	-1.58	-2.58	-1.83	-2.92
$cDCC_2^{Trg}$	–	–	-0.82	-2.36	-0.56	-2.54	-3.79
$cDCC_3^{Trg}$	–	–	–	-1.44	0.19	-3.03	-4.23
$cDCC_5^{Trg}$	–	–	–	–	0.84	-1.87	-3.85
$cDCC_*^{Trg}$	–	–	–	–	–	-1.23	-2.86
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-3.70
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	1.84	1.27	0.88	-0.27	0.39	0.28
$cDCC_2$	–	–	-0.13	0.05	-2.03	-1.09	-0.76
$cDCC_3$	–	–	–	0.19	-1.43	-1.84	-1.05
$cDCC_5$	–	–	–	–	-0.96	-1.66	-2.60
$cDCC_*$	–	–	–	–	–	0.49	0.36
$cDCC_{VC}$	–	–	–	–	–	–	-0.00
$cDCC$	–	–	–	–	–	–	–

$EWMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	1.13	1.46	0.92	-1.27	0.60	-0.29
$cDCC_2^{Trg}$	–	–	1.40	0.04	-2.06	-0.45	-1.52
$cDCC_3^{Trg}$	–	–	–	-1.81	-1.90	-2.49	-2.77
$cDCC_5^{Trg}$	–	–	–	–	-1.53	-0.80	-2.08
$cDCC_*^{Trg}$	–	–	–	–	–	1.13	0.03
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.95
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.43	-0.11	-1.35	-0.15	-1.69	-1.30
$cDCC_2$	–	–	-0.86	-1.80	-0.70	-2.12	-1.82
$cDCC_3$	–	–	–	-2.38	0.02	-2.80	-2.23
$cDCC_5$	–	–	–	–	1.32	-0.96	0.02
$cDCC_*$	–	–	–	–	–	-1.63	-1.32
$cDCC_{VC}$	–	–	–	–	–	–	0.86
$cDCC$	–	–	–	–	–	–	–

$MVMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.53	-1.12	-1.09	-0.51	-0.90	-0.20
$cDCC_2^{Trg}$	–	–	-0.56	-0.99	0.43	-0.58	0.11
$cDCC_3^{Trg}$	–	–	–	-0.29	1.08	0.02	0.48
$cDCC_5^{Trg}$	–	–	–	–	1.01	0.45	1.11
$cDCC_*^{Trg}$	–	–	–	–	–	-0.80	-0.14
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	0.75
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-2.03	-2.39	-1.88	-0.28	-2.05	-1.33
$cDCC_2$	–	–	-0.70	-0.74	1.94	-0.73	0.07
$cDCC_3$	–	–	–	-0.05	2.55	0.13	0.64
$cDCC_5$	–	–	–	–	1.93	0.17	0.85
$cDCC_*$	–	–	–	–	–	-2.14	-1.32
$cDCC_{VC}$	–	–	–	–	–	–	0.60
$cDCC$	–	–	–	–	–	–	–

$MVMSE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-1.07	-0.17	0.05	-0.35	-0.80	-0.39
$cDCC_2^{Trg}$	–	–	0.72	0.74	0.95	0.28	0.67
$cDCC_3^{Trg}$	–	–	–	0.23	0.05	-0.66	-0.24
$cDCC_5^{Trg}$	–	–	–	–	-0.12	-1.22	-0.36
$cDCC_*^{Trg}$	–	–	–	–	–	-0.57	-0.27
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	0.38
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-2.05	-1.76	-1.25	0.25	-2.06	-1.75
$cDCC_2$	–	–	-0.10	0.24	2.12	-0.47	-0.76
$cDCC_3$	–	–	–	0.64	1.89	-0.47	-0.95
$cDCC_5$	–	–	–	–	1.33	-1.02	-1.28
$cDCC_*$	–	–	–	–	–	-2.09	-1.84
$cDCC_{VC}$	–	–	–	–	–	–	-0.59
$cDCC$	–	–	–	–	–	–	–

Table A.2: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Regression-based specification tests

$N = 100$; sequential dynamic parameter estimation.

ENGLE-COLACITO REGRESSION

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	1.94	1.99	1.99	2.05	1.99	2.05	2.27
MV	3.21	3.25	3.60	3.54	3.20	3.51	4.26
MV*	1.50	1.66	1.84	1.85	1.56	1.91	2.64

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	2.37	2.27	2.25	2.28	2.35	2.30	2.27
MV	4.25	4.53	4.65	4.41	4.14	4.44	4.26
MV*	2.72	2.88	2.83	2.68	2.60	2.77	2.64

PORTFOLIO ARCH EFFECT TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	35.60	35.32	35.11	34.80	35.61	35.24	35.77
MV	18.24	16.58	16.59	18.78	19.31	19.45	21.63
MV*	31.15	34.99	25.92	24.40	32.70	25.77	31.62

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	33.01	35.35	34.93	35.06	33.86	35.16	35.77
MV	25.07	23.34	17.26	21.76	26.41	22.61	21.63
MV*	28.50	30.48	26.37	25.40	28.64	26.91	31.62

1% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	1.76	1.80	1.80	1.84	1.81	1.80	1.84
MV	0.90	0.88	0.55	2.02	0.92	0.87	2.02
MV*	1.28	1.66	1.49	1.52	1.29	1.53	2.37

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	1.85	1.85	1.84	1.84	1.85	1.84	1.84
MV	0.47	0.91	1.78	1.61	0.95	2.15	2.02
MV*	1.89	2.51	2.09	2.54	2.33	2.33	2.37

5% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_*^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	0.71	0.71	0.71	0.71	0.72	0.71	0.71
MV	0.97	1.79	0.78	0.68	0.73	0.99	2.36
MV*	1.80	1.36	1.13	1.15	1.15	1.19	1.54

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_*$	$cDCC_{VC}$	$cDCC$
EW	1.10	0.97	0.97	0.97	1.10	0.97	0.71
MV	1.12	1.20	1.12	1.44	1.58	1.98	2.36
MV*	1.32	1.21	1.12	1.55	1.13	1.41	1.54

Table A.3: White cells denote insignificance at 5% level, light grey cells denote significance at 5% level and dark grey cells denote significance at 1% level.

Amisano-Giacomini pairwise comparison tests

$N = 10$; joint dynamic parameter estimation.

$VARSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.40	-1.00	-0.53	-0.37	-0.32	-1.95
$cDCC_2^{Trg}$	–	–	-1.59	-0.61	-0.29	0.30	-1.74
$cDCC_3^{Trg}$	–	–	–	0.18	0.35	2.18	-1.67
$cDCC_5^{Trg}$	–	–	–	–	0.72	0.72	-1.50
$cDCC_8^{Trg}$	–	–	–	–	–	0.37	-1.51
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.81
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.09	-0.12	-0.21	-0.21	-0.12	-1.94
$cDCC_2$	–	–	-0.13	-0.29	-0.29	-0.09	-1.80
$cDCC_3$	–	–	–	-0.37	-0.37	0.09	-1.69
$cDCC_5$	–	–	–	–	-0.18	0.27	-1.26
$cDCC_8$	–	–	–	–	–	0.28	-1.21
$cDCC_{VC}$	–	–	–	–	–	–	-2.03
$cDCC$	–	–	–	–	–	–	–

$CORRSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.53	-2.68	-0.47	-0.51	-0.64	-5.36
$cDCC_2^{Trg}$	–	–	-2.87	-0.16	-0.26	-0.21	-5.39
$cDCC_3^{Trg}$	–	–	–	2.45	2.02	3.09	-5.31
$cDCC_5^{Trg}$	–	–	–	–	-0.30	0.08	-5.56
$cDCC_8^{Trg}$	–	–	–	–	–	0.19	-5.56
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-5.52
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.63	-1.23	-0.83	-0.36	-1.07	-3.58
$cDCC_2$	–	–	-0.96	-0.59	-0.04	-0.80	-3.59
$cDCC_3$	–	–	–	0.35	0.75	0.63	-3.82
$cDCC_5$	–	–	–	–	0.94	0.16	-4.17
$cDCC_8$	–	–	–	–	–	-0.40	-4.26
$cDCC_{VC}$	–	–	–	–	–	–	-3.88
$cDCC$	–	–	–	–	–	–	–

$EWSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.79	0.71	0.51	0.50	0.77	-0.55
$cDCC_2^{Trg}$	–	–	-0.92	-0.10	-0.01	-0.19	-1.37
$cDCC_3^{Trg}$	–	–	–	0.16	0.20	1.01	-1.30
$cDCC_5^{Trg}$	–	–	–	–	0.36	0.07	-1.14
$cDCC_8^{Trg}$	–	–	–	–	–	-0.02	-1.11
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.37
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.83	0.86	0.61	0.58	0.86	0.77
$cDCC_2$	–	–	0.71	0.22	0.20	0.27	-0.58
$cDCC_3$	–	–	–	-0.00	0.02	-0.72	-0.72
$cDCC_5$	–	–	–	–	0.11	-0.16	-0.42
$cDCC_8$	–	–	–	–	–	-0.16	-0.40
$cDCC_{VC}$	–	–	–	–	–	–	-0.67
$cDCC$	–	–	–	–	–	–	–

$MVSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.20	-1.87	-1.96	-2.36	-0.21	-2.60
$cDCC_2^{Trg}$	–	–	-2.06	-2.43	-2.66	-0.71	-2.63
$cDCC_3^{Trg}$	–	–	–	-0.01	-0.76	1.92	-2.55
$cDCC_5^{Trg}$	–	–	–	–	-1.13	2.12	-2.37
$cDCC_8^{Trg}$	–	–	–	–	–	2.38	-2.32
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-2.72
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.29	-1.48	-1.77	-1.33	-0.74	-2.40
$cDCC_2$	–	–	-1.71	-2.03	-1.71	-1.11	-2.47
$cDCC_3$	–	–	–	-0.15	0.43	1.45	-2.47
$cDCC_5$	–	–	–	–	0.79	1.60	-2.38
$cDCC_8$	–	–	–	–	–	1.30	-2.42
$cDCC_{VC}$	–	–	–	–	–	–	-2.73
$cDCC$	–	–	–	–	–	–	–

$MVSCORE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-1.13	-1.97	-2.57	-1.40	-1.17	-2.11
$cDCC_2^{Trg}$	–	–	-1.71	-2.86	-1.24	-0.43	-1.99
$cDCC_3^{Trg}$	–	–	–	-1.31	0.28	1.74	-1.87
$cDCC_5^{Trg}$	–	–	–	–	2.16	2.88	-1.54
$cDCC_8^{Trg}$	–	–	–	–	–	1.13	-1.85
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-2.10
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.92	-1.09	-1.58	-1.22	-1.20	-1.96
$cDCC_2$	–	–	-0.78	-1.47	-1.00	-0.91	-1.82
$cDCC_3$	–	–	–	-1.37	-0.59	0.07	-2.16
$cDCC_5$	–	–	–	–	1.19	1.26	-1.72
$cDCC_8$	–	–	–	–	–	0.57	-1.92
$cDCC_{VC}$	–	–	–	–	–	–	-2.08
$cDCC$	–	–	–	–	–	–	–

Table A.4: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Amisano-Giacomini pairwise comparison tests

$N = 10$; sequential dynamic parameter estimation.

$VARSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.70	-1.42	-1.01	-0.81	-0.63	-1.48
$cDCC_2^{Trg}$	–	–	-1.65	-1.10	-0.71	0.02	-1.45
$cDCC_3^{Trg}$	–	–	–	0.23	0.49	2.17	-1.27
$cDCC_5^{Trg}$	–	–	–	–	0.83	1.18	-1.33
$cDCC_8^{Trg}$	–	–	–	–	–	0.78	-1.39
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.55
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.32	-0.66	-0.67	-0.68	-0.62	-1.00
$cDCC_2$	–	–	-0.82	-0.83	-0.84	-0.65	-1.22
$cDCC_3$	–	–	–	-0.36	-0.49	0.55	-1.49
$cDCC_5$	–	–	–	–	-0.39	0.54	-1.36
$cDCC_8$	–	–	–	–	–	0.60	-1.39
$cDCC_{VC}$	–	–	–	–	–	–	-1.35
$cDCC$	–	–	–	–	–	–	–

$CORRSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-1.17	-3.33	-2.02	-1.85	-1.25	-4.18
$cDCC_2^{Trg}$	–	–	-3.16	-1.80	-1.57	-0.35	-4.23
$cDCC_3^{Trg}$	–	–	–	1.79	1.74	3.29	-3.84
$cDCC_5^{Trg}$	–	–	–	–	0.14	1.75	-4.14
$cDCC_8^{Trg}$	–	–	–	–	–	1.48	-4.17
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-4.39
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.74	-1.98	-1.85	-1.69	-1.52	-2.11
$cDCC_2$	–	–	-1.97	-1.92	-1.75	-1.49	-2.17
$cDCC_3$	–	–	–	0.03	0.11	1.60	-1.39
$cDCC_5$	–	–	–	–	0.19	1.44	-2.01
$cDCC_8$	–	–	–	–	–	1.22	-2.24
$cDCC_{VC}$	–	–	–	–	–	–	-2.20
$cDCC$	–	–	–	–	–	–	–

$EWSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.81	0.74	0.53	0.54	0.79	-0.00
$cDCC_2^{Trg}$	–	–	-0.70	-0.22	-0.09	-0.11	-0.55
$cDCC_3^{Trg}$	–	–	–	0.12	0.19	0.78	-0.43
$cDCC_5^{Trg}$	–	–	–	–	0.54	0.19	-0.61
$cDCC_8^{Trg}$	–	–	–	–	–	0.06	-0.67
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-0.55
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.80	0.87	0.63	0.64	0.82	0.61
$cDCC_2$	–	–	0.77	0.29	0.34	0.17	0.27
$cDCC_3$	–	–	–	-0.01	0.07	-0.93	-0.02
$cDCC_5$	–	–	–	–	0.58	-0.26	-0.05
$cDCC_8$	–	–	–	–	–	-0.31	-0.50
$cDCC_{VC}$	–	–	–	–	–	–	0.23
$cDCC$	–	–	–	–	–	–	–

$MVSCORE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.11	-2.33	-2.41	-2.96	-0.47	-2.48
$cDCC_2^{Trg}$	–	–	-2.34	-2.77	-3.19	-0.87	-2.52
$cDCC_3^{Trg}$	–	–	–	0.73	-0.07	2.13	-2.21
$cDCC_5^{Trg}$	–	–	–	–	-1.60	2.31	-2.19
$cDCC_8^{Trg}$	–	–	–	–	–	2.83	-2.06
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-2.64
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.16	-1.96	-2.20	-2.12	-1.24	-2.28
$cDCC_2$	–	–	-2.09	-2.35	-2.34	-1.51	-2.40
$cDCC_3$	–	–	–	-0.02	-0.13	1.88	-1.64
$cDCC_5$	–	–	–	–	-0.17	2.17	-1.93
$cDCC_8$	–	–	–	–	–	2.81	-1.90
$cDCC_{VC}$	–	–	–	–	–	–	-2.73
$cDCC$	–	–	–	–	–	–	–

$MVSCORE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-1.31	-2.13	-3.09	-2.67	-1.38	-1.81
$cDCC_2^{Trg}$	–	–	-1.76	-3.18	-2.86	-0.72	-1.68
$cDCC_3^{Trg}$	–	–	–	-1.37	-0.80	1.67	-1.38
$cDCC_5^{Trg}$	–	–	–	–	0.75	3.38	-1.04
$cDCC_8^{Trg}$	–	–	–	–	–	2.97	-1.18
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.80
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-1.11	-1.58	-2.06	-2.06	-1.30	-1.77
$cDCC_2$	–	–	-1.38	-1.94	-1.95	-1.04	-1.64
$cDCC_3$	–	–	–	-1.94	-2.57	1.55	-1.83
$cDCC_5$	–	–	–	–	-0.79	3.03	-0.78
$cDCC_8$	–	–	–	–	–	3.15	-0.46
$cDCC_{VC}$	–	–	–	–	–	–	-2.26
$cDCC$	–	–	–	–	–	–	–

Table A.5: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Diebold-Mariano pairwise comparison tests

$N = 10$; joint dynamic parameter estimation.

$VARMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.30	-1.19	-0.81	-0.80	-0.54	-1.55
$cDCC_2^{Trg}$	–	–	-1.00	-0.78	-0.79	-0.26	-1.39
$cDCC_3^{Trg}$	–	–	–	0.58	0.29	1.09	-0.96
$cDCC_5^{Trg}$	–	–	–	–	-0.34	0.59	-1.46
$cDCC_8^{Trg}$	–	–	–	–	–	0.62	-1.39
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.50
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.19	-0.67	-0.76	-0.68	-0.25	-0.47
$cDCC_2$	–	–	-0.60	-0.95	-0.78	-0.02	-0.46
$cDCC_3$	–	–	–	-0.41	-0.25	0.78	-0.04
$cDCC_5$	–	–	–	–	0.18	0.96	0.17
$cDCC_8$	–	–	–	–	–	0.81	0.13
$cDCC_{VC}$	–	–	–	–	–	–	-0.45
$cDCC$	–	–	–	–	–	–	–

$CORRMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.83	0.61	0.90	0.91	0.84	-2.90
$cDCC_2^{Trg}$	–	–	-1.84	0.96	0.97	0.14	-2.20
$cDCC_3^{Trg}$	–	–	–	1.11	1.09	2.11	-2.20
$cDCC_5^{Trg}$	–	–	–	–	0.98	-0.95	-1.86
$cDCC_8^{Trg}$	–	–	–	–	–	-0.96	-1.80
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-2.21
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.87	0.92	0.90	0.90	0.86	-2.69
$cDCC_2$	–	–	1.00	0.93	0.92	-0.13	-1.55
$cDCC_3$	–	–	–	0.85	0.86	-1.03	-1.40
$cDCC_5$	–	–	–	–	0.85	-0.93	-1.25
$cDCC_8$	–	–	–	–	–	-0.93	-1.23
$cDCC_{VC}$	–	–	–	–	–	–	-1.56
$cDCC$	–	–	–	–	–	–	–

$EWMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.85	0.19	-0.42	-0.32	0.72	-1.28
$cDCC_2^{Trg}$	–	–	-0.87	-1.52	-1.43	-0.20	-1.48
$cDCC_3^{Trg}$	–	–	–	-1.44	-1.21	0.96	-1.44
$cDCC_5^{Trg}$	–	–	–	–	0.53	1.44	-1.27
$cDCC_8^{Trg}$	–	–	–	–	–	1.42	-1.32
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.53
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.53	0.74	-0.27	-0.27	0.66	0.03
$cDCC_2$	–	–	0.69	-0.79	-0.80	0.49	-0.42
$cDCC_3$	–	–	–	-1.39	-1.47	-0.44	-1.15
$cDCC_5$	–	–	–	–	0.09	0.95	0.54
$cDCC_8$	–	–	–	–	–	0.99	0.52
$cDCC_{VC}$	–	–	–	–	–	–	-0.78
$cDCC$	–	–	–	–	–	–	–

$MVMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.48	-1.32	-0.82	-1.91	-0.09	-1.67
$cDCC_2^{Trg}$	–	–	-1.61	-1.73	-1.69	-0.98	-1.89
$cDCC_3^{Trg}$	–	–	–	0.58	-0.69	1.32	-1.63
$cDCC_5^{Trg}$	–	–	–	–	-1.07	0.84	-1.66
$cDCC_8^{Trg}$	–	–	–	–	–	1.29	-1.36
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.96
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.54	-1.36	-1.52	-0.66	-0.72	-1.45
$cDCC_2$	–	–	-1.68	-1.92	-1.26	-1.17	-1.75
$cDCC_3$	–	–	–	0.48	1.16	1.68	-1.20
$cDCC_5$	–	–	–	–	1.19	1.08	-1.26
$cDCC_8$	–	–	–	–	–	-0.32	-1.87
$cDCC_{VC}$	–	–	–	–	–	–	-2.22
$cDCC$	–	–	–	–	–	–	–

$MVMSE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.10	-0.60	-0.97	0.17	-0.31	-1.63
$cDCC_2^{Trg}$	–	–	-1.28	-1.50	0.22	-1.10	-1.76
$cDCC_3^{Trg}$	–	–	–	-0.25	1.46	0.83	-1.63
$cDCC_5^{Trg}$	–	–	–	–	1.06	1.06	-1.47
$cDCC_8^{Trg}$	–	–	–	–	–	-0.70	-1.96
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.79
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.02	-1.15	-0.97	-0.11	-0.63	-1.86
$cDCC_2$	–	–	-1.42	-1.11	-0.18	-0.92	-1.94
$cDCC_3$	–	–	–	0.44	1.57	1.51	-2.17
$cDCC_5$	–	–	–	–	0.99	0.56	-2.10
$cDCC_8$	–	–	–	–	–	-0.89	-2.43
$cDCC_{VC}$	–	–	–	–	–	–	-2.33
$cDCC$	–	–	–	–	–	–	–

Table A.6: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Diebold-Mariano pairwise comparison tests

$N = 10$; sequential dynamic parameter estimation.

$VARMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.27	-1.35	-0.83	-0.92	-0.79	-1.72
$cDCC_2^{Trg}$	–	–	-1.38	-0.72	-0.91	-0.59	-1.47
$cDCC_3^{Trg}$	–	–	–	0.93	0.55	1.43	-0.83
$cDCC_5^{Trg}$	–	–	–	–	-0.96	0.18	-1.58
$cDCC_8^{Trg}$	–	–	–	–	–	0.45	-1.33
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.58
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.21	-1.09	-0.85	-0.85	-0.58	-0.88
$cDCC_2$	–	–	-1.19	-0.86	-0.90	-0.43	-0.93
$cDCC_3$	–	–	–	0.60	0.43	1.29	0.33
$cDCC_5$	–	–	–	–	-0.43	0.50	-0.65
$cDCC_8$	–	–	–	–	–	0.59	-0.54
$cDCC_{VC}$	–	–	–	–	–	–	-0.66
$cDCC$	–	–	–	–	–	–	–

$CORRMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.72	0.31	0.66	0.69	0.72	-0.70
$cDCC_2^{Trg}$	–	–	-1.81	0.60	0.67	0.12	-1.50
$cDCC_3^{Trg}$	–	–	–	0.92	0.93	2.20	-1.17
$cDCC_5^{Trg}$	–	–	–	–	0.92	-0.59	-2.71
$cDCC_8^{Trg}$	–	–	–	–	–	-0.66	-2.91
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.54
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.85	0.83	0.80	0.80	0.74	0.65
$cDCC_2$	–	–	0.79	0.76	0.76	-0.69	0.43
$cDCC_3$	–	–	–	0.70	0.71	-0.96	-0.43
$cDCC_5$	–	–	–	–	0.52	-0.85	-1.55
$cDCC_8$	–	–	–	–	–	-0.84	-1.50
$cDCC_{VC}$	–	–	–	–	–	–	0.55
$cDCC$	–	–	–	–	–	–	–

$EWMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.84	0.07	-0.46	-0.34	0.46	-1.23
$cDCC_2^{Trg}$	–	–	-0.83	-1.63	-1.47	-0.50	-1.42
$cDCC_3^{Trg}$	–	–	–	-1.40	-1.06	0.69	-1.40
$cDCC_5^{Trg}$	–	–	–	–	0.65	1.33	-1.24
$cDCC_8^{Trg}$	–	–	–	–	–	1.26	-1.28
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.46
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.47	0.31	-0.41	-0.33	0.25	-0.50
$cDCC_2$	–	–	-0.01	-0.95	-0.87	-0.21	-1.02
$cDCC_3$	–	–	–	-1.52	-1.62	-0.27	-1.97
$cDCC_5$	–	–	–	–	0.74	0.89	-0.19
$cDCC_8$	–	–	–	–	–	0.86	-0.76
$cDCC_{VC}$	–	–	–	–	–	–	-1.18
$cDCC$	–	–	–	–	–	–	–

$MVMSE$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	0.36	-1.75	-0.61	-2.20	-0.33	-1.65
$cDCC_2^{Trg}$	–	–	-1.86	-1.79	-1.97	-1.05	-1.91
$cDCC_3^{Trg}$	–	–	–	1.32	0.55	1.57	-1.39
$cDCC_5^{Trg}$	–	–	–	–	-1.42	0.15	-1.69
$cDCC_8^{Trg}$	–	–	–	–	–	1.17	-1.34
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-2.06
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	0.37	-1.52	-1.53	-1.17	-0.94	-1.42
$cDCC_2$	–	–	-1.77	-1.89	-1.59	-1.27	-1.79
$cDCC_3$	–	–	–	1.09	1.10	1.99	0.49
$cDCC_5$	–	–	–	–	0.32	1.07	-0.72
$cDCC_8$	–	–	–	–	–	1.11	-1.56
$cDCC_{VC}$	–	–	–	–	–	–	-2.17
$cDCC$	–	–	–	–	–	–	–

$MVMSE^*$

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
$cDCC_1^{Trg}$	–	-0.03	-0.96	-1.19	-1.04	-0.46	-1.46
$cDCC_2^{Trg}$	–	–	-1.43	-2.01	-1.86	-0.88	-1.54
$cDCC_3^{Trg}$	–	–	–	0.02	0.22	1.15	-1.16
$cDCC_5^{Trg}$	–	–	–	–	0.29	1.89	-1.16
$cDCC_8^{Trg}$	–	–	–	–	–	1.65	-1.22
$cDCC_{VC}^{Trg}$	–	–	–	–	–	–	-1.61
$cDCC$	–	–	–	–	–	–	–

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
$cDCC_1$	–	-0.12	-1.70	-1.65	-1.51	-0.95	-1.74
$cDCC_2$	–	–	-1.80	-1.74	-1.60	-1.04	-1.77
$cDCC_3$	–	–	–	0.75	0.24	2.63	-0.55
$cDCC_5$	–	–	–	–	-0.46	1.89	-1.48
$cDCC_8$	–	–	–	–	–	2.37	-0.89
$cDCC_{VC}$	–	–	–	–	–	–	-2.33
$cDCC$	–	–	–	–	–	–	–

Table A.7: Negative (resp., positive) values provide evidence in favor of the model in row (resp., column); white cells denote equal predictive ability at 5% level; grey cells denote rejection of the null of equal predictive ability at 5% level.

Regression-based specification tests

$N = 10$; joint dynamic parameter estimation.

ENGLE-COLACITO REGRESSION

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	0.64	0.62	0.65	0.61	0.60	0.63	1.00
MV	1.91	2.04	2.23	2.02	2.15	2.04	2.91
MV*	0.80	0.90	1.00	0.93	0.88	0.91	1.57

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	0.93	0.92	0.92	0.91	0.90	0.93	1.00
MV	2.10	2.23	2.25	2.38	2.43	2.27	2.91
MV*	1.09	1.21	1.22	1.27	1.26	1.26	1.57

PORTFOLIO ARCH EFFECT TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	10.06	11.04	11.10	11.57	11.66	11.09	11.98
MV	10.23	10.22	12.71	13.46	12.36	10.74	48.12
MV*	12.59	12.78	13.44	12.72	13.24	12.67	19.24

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	11.33	12.42	12.67	13.19	13.35	12.46	11.98
MV	13.56	14.66	27.68	26.73	26.21	20.63	48.12
MV*	14.92	15.15	17.54	16.35	16.69	15.53	19.24

1% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	1.25	1.25	1.24	1.25	1.24	1.25	1.44
MV	0.65	0.77	1.43	0.97	0.87	1.59	2.36
MV*	1.29	1.94	1.05	1.26	1.05	1.94	1.72

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	1.44	1.44	1.44	1.44	1.44	1.44	1.44
MV	0.94	0.87	2.29	2.07	1.75	1.72	2.36
MV*	1.83	1.32	1.57	1.86	1.86	1.33	1.72

5% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	2.19	2.18	2.18	2.19	2.19	2.18	2.45
MV	0.97	1.03	0.90	1.82	1.22	1.03	1.41
MV*	1.56	1.56	1.77	1.78	1.78	1.77	2.41

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	2.46	2.45	2.45	2.45	2.45	2.45	2.45
MV	1.30	1.33	1.30	1.32	1.06	1.36	1.41
MV*	2.18	2.46	2.20	1.96	1.97	2.47	2.41

Table A.8: White cells denote insignificance at 5% level, light grey cells denote significance at 5% level and dark grey cells denote significance at 1% level.

Regression-based specification tests

$N = 10$; sequential dynamic parameter estimation.

ENGLE-COLACITO REGRESSION

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	0.73	0.73	0.76	0.75	0.75	0.73	1.00
MV	2.21	2.34	2.62	2.44	2.59	2.36	2.91
MV*	0.98	1.10	1.27	1.22	1.19	1.13	1.57

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	0.98	0.98	0.99	0.99	0.98	0.99	1.00
MV	2.33	2.45	2.67	2.81	2.89	2.56	2.91
MV*	1.23	1.35	1.50	1.57	1.58	1.46	1.57

PORTFOLIO ARCH EFFECT TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	9.05	9.92	10.00	10.37	10.39	9.97	11.98
MV	9.07	9.39	13.51	14.55	12.50	10.09	48.12
MV*	11.24	11.46	12.50	11.41	11.60	11.28	19.24

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	10.09	11.13	11.39	11.82	11.87	11.19	11.98
MV	11.79	13.76	38.30	37.27	37.96	28.24	48.12
MV*	13.33	13.76	17.47	16.09	16.97	14.88	19.24

1% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	1.25	1.25	1.24	1.24	1.24	1.24	1.44
MV	0.64	0.76	2.03	1.17	1.36	1.60	2.36
MV*	1.29	1.75	1.32	1.05	0.89	1.94	1.72

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	1.45	1.18	1.44	1.44	1.44	1.44	1.44
MV	0.81	0.75	1.88	1.84	1.90	1.52	2.36
MV*	1.55	1.55	1.45	1.87	1.70	1.57	1.72

5% DYNAMIC QUANTILE TEST

	$cDCC_1^{Trg}$	$cDCC_2^{Trg}$	$cDCC_3^{Trg}$	$cDCC_5^{Trg}$	$cDCC_8^{Trg}$	$cDCC_{VC}^{Trg}$	$cDCC$
EW	2.19	2.18	2.18	2.19	2.18	2.18	2.45
MV	0.73	1.09	0.66	1.14	1.25	1.11	1.41
MV*	1.55	1.55	1.77	1.77	1.77	1.77	2.41

	$cDCC_1$	$cDCC_2$	$cDCC_3$	$cDCC_5$	$cDCC_8$	$cDCC_{VC}$	$cDCC$
EW	2.45	2.45	2.44	2.45	2.45	2.44	2.45
MV	1.16	1.18	1.22	0.81	1.05	1.26	1.41
MV*	2.18	2.46	2.39	1.96	2.14	2.48	2.41

Table A.9: White cells denote insignificance at 5% level, light grey cells denote significance at 5% level and dark grey cells denote significance at 1% level.

Datasets

	LARGE DATASET ($N = 100$)		SMALL DATASET ($N = 10$)
	Industrial equities	Consumer good equities	S&P 500 Index and related sector SPDR's
1	BRUSH ENGD.MATERIALS	WD-40	CONSUMER DISCRETIONARY SPDR
2	FISERV	LEGGETT&PLATT	CONSUMER STAPLES SPDR
3	CON-WAY	GREEN MNT.COF.ROASTERS	ENERGY SPDR
4	CLARCOR	M/I HOMES	FINANCIALS SPDR
5	COMMSCOPE	THE HERSHEY COMPANY	HEALTCARE SPDR
6	AUTOMATIC DATA PROC.	LANCE	INDUSTRIALS SPDR
7	NATIONAL INSTS.	CALLAWAY GOLF	MATERIALS SPDR
8	ON ASSIGNMENT	RC2	TECHNOLOGY SPDR
9	CYBERSOURCE	HNI	UTILITIES SPDR
10	SFN GROUP	POLARIS INDS.	S&P 500 COMPOSITE
11	FTI CONSULTING	MDC HDG.	
12	DELUXE	DECKERS OUTDOOR	
13	ROBERT HALF INTL.	FORD MOTOR	
14	IRON MNT.	HJ HEINZ	
15	JACOBS ENGR.	K-SWISS 'A'	
16	NEWPORT	CHURCH & DWIGHT CO.	
17	ROCK-TENN 'A' SHS.	COLGATE-PALM.	
18	STERICYCLE	KIMBERLY-CLARK	
19	LINCOLN ELECTRIC HDG.	HARLEY-DAVIDSON	
20	CATERPILLAR	BRUNSWICK	
21	BELDEN	SCOTTS MIRACLE-GRO	
22	G & K SVS.'A'	SARA LEE	
23	MOOG 'A'	TIMBERLAND 'A'	
24	CUMMINS	BRIGGS & STRATTON	
25	GEO GROUP	TYSON FOODS 'A'	
26	PARKER-HANNIFIN	CLOROX	
27	SYKES ENTERPRISES	THOR INDUSTRIES	
28	HUB GROUP 'A'	GOODYEAR TIRE & RUB.	
29	TELEDYNE TECHS.	REYNOLDS AMERICAN	
30	ALEX.& BALDWIN	ESTEE LAUDER COS.'A'	
31	KNIGHT TRANSPORTATION	TOLL BROS.	
32	CINTAS	KID BRANDS	
33	PERKINELMER	LA-Z-BOY	
34	WW GRAINGER	MANNATECH	
35	BEMIS	CORN PRODUCTS INTL.	
36	PREC.CASTPARTS	NIKE 'B'	
37	CURTISS WRIGHT	ARCTIC CAT	
38	EMCOR GROUP	LIZ CLAIBORNE	
39	PARK ELECTROCHEMICAL	ELECTRONIC ARTS	
40	ARROW ELECTRONICS	SMITHFIELD FOODS	
41	BALDOR ELECTRIC	ETHAN ALLEN INTERIORS	
42	TECHNITROL	SCHWEITZER-MAUDUIT INT.	
43	CARLISLE COS.	COCA COLA ENTS.	
44	KIRBY	DREW INDS.	
45	LAWSON PRODUCTS	FOSSIL	
46	COMFORT SYS.USA	TOOTSIE ROLL	
47	MEADWESTVACO	ARCHER-DANLS.-MIDL.	
48	KAMAN 'A'		
49	ENCORE WIRE		
50	CLEAN HARBORS		
51	HEARTLAND EXPRESS		
52	CUBIC		
53	ACTUANT 'A'		

Table A.10: Assets included in the large dataset and in the small dataset. The assets in the large datasets are randomly selected equities from the industrial and consumer good components of the S&P 1500 Index.